Behavioral Game Theory: Thinking, Learning, and Teaching

Colin F. Camerer[1]
California Institute of Technology
Pasadena, CA 91125

Teck-Hua Ho
Wharton School, University of Pennsylvania
Philadelphia PA 19104

Juin Kuan Chong
National University of Singapore
Kent Ridge Crescent
Singapore 119260

November 14, 2001

[1]This research was supported by NSF grants SBR 9730364, SBR 9730187 and SES-0078911. Thanks to many people for helpful comments on this research, particularly Caltech colleagues (especially Richard McKelvey, Tom Palfrey, and Charles Plott), Mónica Capra, Vince Crawford, John Duffy, Drew Fudenberg, John Kagel, members of the MacArthur Preferences Network, our research assistants and collaborators Dan Clendenning, Graham Free, David Hsia, Ming Hsu, Hongjai Rhee, and Xin Wang, and seminar audience members too numerous to mention. Dan Levin gave the shooting-ahead military example. Dave Cooper, Ido Erev, and Bill Frechette wrote helpful emails.

VISIT…

1 Introduction

Game theory is a mathematical system for analyzing and predicting how humans behave in strategic situations. Standard equilibrium analyses assume all players: 1) form beliefs based on analysis of what others might do (strategic thinking); 2) choose a best response given those beliefs (optimization); 3) adjust best responses and beliefs until they are mutually consistent (equilibrium).

It is widely-accepted that not every player behaves rationally in complex situations, so assumptions (1) and (2) are sometimes violated. For explaining consumer choices and other decisions, rationality may still be an adequate approximation even if a modest percentage of players violate the theory. But game theory is different. Players' fates are intertwined. The presence of players who do not think strategically or optimize can therefore change what rational players should do. As a result, what a population of players is likely to do when some are not thinking strategically and optimizing can only be predicted by an analysis which uses the tools of (1)-(3) but accounts for bounded rationality as well, preferably in a precise way.[2]

It is also unlikely that equilibrium (3) is reached instantaneously in one-shot games. The idea of instant equilibration is so unnatural that perhaps an equilibrium should not be thought of as a prediction which is vulnerable to falsification at all. Instead, it should be thought of as the limiting outcome of an unspecified learning or evolutionary process that unfolds over time.[3] In this view, equilibrium is the *end* of the story of how strategic thinking, optimization, and equilibration (or learning) work, not the beginning (one-shot) or the middle (equilibration).

This paper has three goals. First we develop an index of bounded rationality which measures players' steps of thinking and uses one parameter to specify how heterogeneous a population of players is. Coupled with best response, this index makes a unique prediction of behavior in any one-shot game. Second, we develop a learning algorithm (called Functional Experience-Weighted Attraction Learning (fEWA)) to compute the path of

[2]Our models are related to important concepts like rationalizability, which weakens the mutual consistency requirement, and behavior of finite automata. The difference is that we work with simple parametric forms and concentrate on fitting them to data.

[3]In his thesis proposing a concept of equilibrium, Nash himself suggested equilibrium might arise from some "mass action" which adapted over time. Taking up Nash's implicit suggestion, later analyses filled in details of where evolutionary dynamics lead (see Weibull, 1995; Mailath, 1998).

equilibration. The algorithm generalizes both fictitious play and reinforcement models and has shown greater empirical predictive power than those models in many games (adjusting for complexity, of course). Consequently, fEWA can serve as an empirical device for finding the behavioral resting point as a function of the initial conditions. Third, we show how the index of bounded rationality and the learning algorithm can be used to understand repeated game behaviors such as reputation building and strategic teaching.

Our approach is guided by three stylistic principles: Precision; generality; and empirical discipline. The first two are standard desiderata in game theory; the third is a cornerstone in experimental economics.

Precision: Because game theory predictions are sharp, it is not hard to spot likely deviations and counterexamples. Until recently, most of the experimental literature consisted of documenting deviations (or successes) and presenting a simple model, usually specialized to the game at hand. The hard part is to distill the deviations into an alternative theory that is similarly precise as standard theory and can be widely applied. We favor specifications that use one or two free parameters to express crucial elements of behavioral flexibility because people are different. We also prefer to let data, rather than our intuition, specify parameter values.[4]

Generality: Much of the power of equilibrium analyses, and their widespread use, comes from the fact that the same principles can be applied to many different games, using the universal language of mathematics. Widespread use of the language creates a dialogue that sharpens theory and cumulates worldwide knowhow. Behavioral models of games are also meant to be general, in the sense that the same models can be applied to many games with minimal customization. The insistence on generality is common in economics, but is not universal. Many researchers in psychology believe that behavior is so context-specific that it is impossible to have a common theory that applies to all contexts. Our view is that we can't know whether general theories fail until they are broadly applied. Showing that customized models of different games fit well does not mean there isn't a general theory waiting to be discovered that is even better.

[4]While great triumphs of economic theory come from parameter-free models (e.g., Nash equilibrium), relying on a small number of free parameters is more typical in economic modeling. For example, nothing in the theory of intertemporal choice pins a discount factor δ to a specific value. But if a wide range of phenomena are consistent with a value like .95, then as economists we are comfortable working with such a value despite the fact that it does not emerge from axioms or deeper principles.

It is noteworthy that in the search for generality, the models we describe below are typically fit to dozens of different data sets, rather than one or two. The number of subject-periods used when games are pooled is usually several thousand. This doesn't mean the results are conclusive or unshakeable. It just illustrates what we mean by a general model.

Empirical discipline: Our approach is heavily disciplined by data. Because game theory is about people (and groups of people) thinking about what other people and groups will do, it is unlikely that pure logic alone will tell us what they will happen.[5] As the physicist Murray Gell-Mann said, 'Think how hard physics would be if particles could think.' It is even harder if we don't watch what 'particles' do when interacting.

Our insistence on empirical discipline is shared by others, past and present. Von Neumann and Morgenstern (1944) thought that

> the empirical background of economic science is definitely inadequate...it would have been absurd in physics to expect Kepler and Newton without Tycho Brahe,– and there is no reason to hope for an easier development in economics

Fifty years later Eric Van Damme (1999) thought the same:

> Without having a broad set of facts on which to theorize, there is a certain danger of spending too much time on models that are mathematically elegant, yet have little connection to actual behavior. At present our empirical knowledge is inadequate and it is an interesting question why game theorists have not turned more frequently to psychologists for information about the learning and information processes used by humans.

The data we use to inform theory are experimental because game-theoretic predictions are notoriously sensitive to what players know, when they move, and what their payoffs are. Laboratory environments provide crucial control of all these variables (see Crawford, 1997). As in other lab sciences, the idea is to use lab control to sort out which theories

[5] As Thomas Schelling (1960, p. 164) wrote "One cannot, without empirical evidence, deduce what understandings can be perceived in a nonzero-sum game of maneuver any more than one can prove, by purely formal deduction, that a particular joke is bound to be funny."

work well and which don't, then later use them to help understand patterns in naturally-occurring data. In this respect, behavioral game theory resembles data-driven fields like labor economics or finance more than analytical game theory. The large body of experimental data accumulated over the last couple of decades (and particularly the last five years; see Camerer, 2002) is a treasure trove which can be used to sort out which simple parametric models fit well.

While the primary goal of behavioral game theory models is to make accurate predictions when equilibrium concepts do not, it can also circumvent two central problems in game theory: Refinement and selection. Because we replace the strict best-response (optimization) assumption with stochastic better-response, all possible paths are part of a (statistical) equilibrium. As a result, there is no need to apply subgame perfection or propose belief refinements (to update beliefs after zero-probability events where Bayes' rule is helpless). Furthermore, with plausible parameter values the thinking and learning models often solve the long-standing problem of selecting one of several Nash equilibria, in a statistical sense, because the models make a unimodal statistical prediction rather than predicting multiple modes. Therefore, while the thinking-steps model generalizes the concept of equilibrium, it can also be *more* precise (in a statistical sense) when equilibrium is imprecise (cf. Lucas, 1986).[6]

We make three remarks before proceeding. First, while we do believe the thinking, learning and teaching models in this paper do a good job of explaining some experimental regularity parsimoniously, lots of other models are being actively explored.[7] The models in this paper illustrate what most other models also strive to explain, and how they are

[6]Lucas (1986) makes a similar point in macroeconomic models. Rational expectations often yields indeterminacy whereas adaptive expectations pins down a dynamic path. Lucas writes (p. S421): "The issue involves a question concerning how collections of people behave in a specific situation. Economic theory does not resolve the question...It is hard to see what can advance the discussion short of assembling a collection of people, putting them in the situation of interest, and observing what they do."

[7]Quantal response equilibrium (QRE), a statistical generalization of Nash, almost always explains the direction of deviations from Nash and should replace Nash as the static benchmark that other models are routinely compared to (see Goeree and Holt, in press. Stahl and Wilson (1995), Capra (1999) and Goeree and Holt (1999b) have models of limited thinking in one-shot games which are similar to ours. There are many learning models. fEWA generalizes some of them (though reinforcement with payoff variability adjustment is different; see Erev, Bereby-Meyer, and Roth, 1999). Other approaches include rule learning (Stahl, 1996, 2000), and earlier AI tools like genetic algorithms or genetic programming to "breed" rules. Finally, there are no alternative models of strategic teaching that we know of but this is an important area others should look at.

evaluated.

The second remark is that these behavioral models are shaped by data from game experiments, but are intended for eventual use in areas of economics where game theory has been applied successfully. We will return to a list of potential applications in the conclusion, but to whet the reader's appetite here is a preview. Limited thinking models might be useful in explaining price bubbles, speculation and betting, competition neglect in business strategy, simplicity of incentive contracts, and persistence of nominal shocks in macroeconomics. Learning might be helpful for explaining evolution of pricing, institutions and industry structure. Teaching can be applied to repeated contracting, industrial organization, trust-building, and policymakers setting inflation rates.

The third remark is about how to read this long paper. The second and third sections, on learning and teaching, are based on published research and an unpublished paper introducing the one-parameter functional (fEWA) approach. The first section, on thinking, is new and more tentative. We put all three in one paper to show the ambitions of behavioral game theory– to explain observed regularity in many different games with only a few parameters that codify behavioral intuitions and principles.

2 A thinking model and bounded rationality measure

The thinking model is designed to predict behavior in one-shot games and also to provide initial conditions for models of learning.

We begin with notation. Strategies have numerical attractions that determine the probabilities of choosing different strategies through a logistic response function. For player i, there are m_i strategies (indexed by j) which have initial attractions denoted $A_i^j(0)$. Denote i's jth strategy by s_i^j, chosen strategies by i and other players (denoted $-i$) in period t as $s_i(t)$ and $s_{-i}(t)$, and player i's payoffs of choosing s_i^j by $\pi_i(s_i^j, s_{-i}(t))$.

A logit response rule is used to map attractions into probabilities:

$$P_i^j(t+1) = \frac{e^{\lambda \cdot A_i^j(t)}}{\sum_{k=1}^{m_i} e^{\lambda \cdot A_i^k(t)}} \tag{2.1}$$

where λ is the response sensitivity.[8]

We model thinking by characterizing the number of steps of iterated thinking that subjects do, and their decision rules.[9] In the thinking-steps model some players, using zero steps of thinking, do not reason strategically at all. (Think of these players as being fatigued, clueless, overwhelmed, uncooperative, or simply more willing to make a random guess in the first period of a game and learn from subsequent experience than to think hard before learning.) We assume that zero-step players randomize equally over all strategies.

Players who do one step of thinking *do* reason strategically. What exactly do they do? We assume they are "overconfident"– though they use one step, they believe others are all using zero steps. Proceeding inductively, players who use K steps think all others use zero to $K-1$ steps.

It is useful to ask why the number of steps of thinking might be limited. One answer comes from psychology. Steps of thinking strain "working memory", where items are stored while being processed. Loosely speaking, working memory is a hard constraint. For example, most people can remember only about 5-9 digits when shown a long list of digits (though there are reliable individual differences, correlated with reasoning ability). The strategic question "If she thinks he anticipates what she will do what should she do?" is an example of a recursive "embedded sentence" of the sort that is known to strain working memory and produce inference and recall mistakes.[10]

Reasoning about others might also be limited because players are not certain about another player's payoffs or degree of rationality. Why should they be? After all, adherence to optimization and instant equilibration is a matter of personal taste or skill. But whether other players do the same is a guess about the world (and iterating further, a guess about the contents of another player's brain or a firm's boardroom activity).

[8]Note the timing convention– attractions are defined *before* a period of play; so the initial attractions $A_i^j(0)$ determine choices in period 1, and so forth.

[9]This concept was first studied by Stahl and Wilson (1995) and Nagel (1995), and later by Ho, Camerer and Weigelt (1998). See also Sonsino, Erev and Gilat (2000).

[10]Embedded sentences are those in which subject-object clauses are separated by other subject-object clauses. A classic example is "The mouse that the cat that the dog chased bit ran away". To answer the question "Who got bit?" the reader must keep in mind "the mouse" while processing the fact that the cat was chased by the dog. Limited working memory leads to frequent mistakes in recalling the contents of such sentences or answering questions about them (Christiansen and Chater, 1999). This notation makes it easier: "The mouse that [the cat that [the dog {chased}] bit] ran away".

The key challenge in thinking steps models is pinning down the frequencies of players using different numbers of thinking steps. We assume those frequencies have a Poisson distribution with mean and standard deviation τ (the frequency of level K types is $f(K) = \frac{e^{-\tau}\tau^K}{K!}$). Then τ is an index of bounded rationality.

The Poisson distribution has three appealing properties: It has only one free parameter (τ); since Poisson is discrete it generates "spikes" in predicted distributions reflecting individual heterogeneity (other approaches do not[11]); and for sensible τ values the frequency of step types is similar to the frequencies estimated in earlier studies (see Stahl and Wilson (1995); Ho, Camerer and Weigelt (1998); and Nagel et al., 1999). Figure 1 shows four Poisson distributions with different τ values. Note that there are substantial frequencies of steps 0-3 for τ around one or two. There are also very few higher-step types, which is plausible if the limit on working memory has an upper bound.

Modeling heterogeneity is important because it allows the possibility that not every player is rational. The few studies that have looked carefully found fairly reliable individual differences, because a subject's step level or decision rule is fairly stable across games (Stahl and Wilson, 1995; Costa-Gomes et al., 2001). Including heterogeneity can also improve learning models by starting them off with enough persistent variation across people to match the variation we see across actual people.

To make the model precise, assume players know the absolute frequencies of players at lower levels from the Poisson distribution. But since they do not imagine higher-step types there is missing probability. They must adjust their beliefs by allocating the missing probability in order to compute sensible expected payoffs to guide choices. We assume players divide the correct relative proportions of lower-step types by $\sum_{c=1}^{K-1} f(c)$

[11] A natural competitor to the thinking-steps model for explaining one-shot games is quantal response equilibrium (QRE; see McKelvey and Palfrey, 1995, 1998; Goeree and Holt, 1999a). Weiszacker (2000) suggests an asymmetric version which is equivalent to a thinking steps model in which one type thinks others are more random than she is. More cognitive alternatives are the theory of thinking trees due to Capra (1999) and the theory of "noisy introspection" due to Goeree and Holt (1999b). In Capra's model players introspect until their choices match those of players whose choices they anticipate. In Goeree and Holt's theory players use an iterated quantal response function with a response sensitivity parameter equal to λ/t^n where n is the discrete iteration step. When t is very large, their model corresponds to one in which all players do one step and think others do zero. When $t = 1$ the model is QRE. All these models generate unimodal distributions so they need to be expanded to accommodate heterogeneity. Further work should try to distinguish different models or investigate whether they are similar enough to be close modeling substitutes.

so the adjusted frequencies maintain the same relative proportions but add up to one.

Given this assumption, players using $K > 0$ steps are assumed to compute expected payoffs given their adjusted beliefs, and use those attractions to determine choice probabilities according to

$$A_i^j(0|K) = \sum_{h=1}^{m_{-i}} \pi_i(s_i^j, s_{-i}^h) \cdot \{ \sum_{c=0}^{K-1} [\frac{f(c)}{\sum_{c=0}^{K-1} f(c)} \cdot P_{-i}^h(1|c)]\} \tag{2.2}$$

where $A_i^j(0|K)$ and $P_i^l(1|c))$ are the attraction of level K in period 0 and the predicted choice probability of lower level c in period 1.

As a benchmark we also fit quantal response equilibrium (QRE), defined by

$$A_i^j(0|K) = \sum_{h=1}^{m_{-i}} \pi_i(s_i^j, s_{-i}^h) \cdot P_{-i}^h(1) \tag{2.3}$$

$$P_i^j(1) = \frac{e^{\lambda \cdot A_i^j(0)}}{\sum_{h=1}^{m_i} e^{\lambda \cdot A_i^h(0)}} \tag{2.4}$$

When λ goes to infinity QRE converges to Nash equilibrium. QRE is closely related to a thinking-steps model in which K-step types are "self-aware" and believe there are other K-step types, and τ goes to infinity.

2.1 Fitting the model

As a first pass the thinking-steps model was fit to data from three studies in which players made decisions in matrix games once each without feedback (a total of 2558 subject-games).[12] Within each of the three data sets, a common λ was used, and best-fitting τ values were estimated both separately for each game, and fixed across games (maximizing log likelihood).

Table 1 reports τ values for each game separately, common τ and λ from the thinking steps model, and measures of fit for the thinking model and QRE– the log likelihood LL (which can be used to compare models) and the mean of the squared deviations (MSD) between predicted and actual frequencies.

[12]The data are 48 subjects playing 12 symmetric 3x3 games (Stahl and Wilson, 1995), 187 subjects playing 8 2x2 asymmetric matrix games (Cooper and Van Huyck, 2001) and 36 subjects playing 13 asymmetric games ranging from 2x2 to 4x2 (Costa-Gomes, Crawford and Broseta, 2001).

Table 1: Estimates of thinking model τ and fit statistics, 3 matrix game experiments

	Stahl and Wilson (1995a)	Cooper Van Huyck (2001)	Costa-Gomes et al. (2001)
	game-specific τ estimates		
Game 1	18.34	1.14	2.17
Game 2	2.26	1.04	2.21
Game 3	1.99	0.00	2.22
Game 4	4.56	1.25	1.44
Game 5	5.53	0.53	1.81
Game 6	1.70	0.80	1.58
Game 7	5.55	1.17	1.08
Game 8	2.03	1.75	1.94
Game 9	1.79		1.88
Game 10	8.79		2.66
Game 11	7.33		1.34
Game 12	21.46		2.30
Game 13			2.36
common τ	8.44	0.81	2.22
common λ	9.06	190.58	15.76
	fit statistics (thinking steps model)		
MSD (pooled)	0.0257	0.0135	0.0063
LL (pooled)	-1115	-1739	-555
	fit statistics (QRE)		
MSD (QRE)	0.0327	0.0269	0.0079
LL (QRE)	-1176	-1838	-599

Note: In Costa-Gomes et al. the games are labeled as 2b 2x2,3a 2x2, 3b 2x2, 4b 3x2, 4c 3x2, 5b 3x2, 8b 3x2, 9a 4x2, 4a 2x3, 4d 2x3, 6b 2x3, 7b 2x3, 9b 2x4.

QRE fits a little worse than the thinking model in all three data sets.[13] This is a big clue that an overconfidence specification is more realistic than one with self-awareness.

Estimated values of τ are quite variable in the Stahl and Wilson data but fairly consistent in the others.[14] In the latter two sets of data, estimates are clustered around one and two, respectively. Imposing a common τ across games only reduces fit very slightly (even in the Stahl and Wilson game[15].) The fact that the cross-game estimates are the most consistent in the Costa-Gomes et al. games, which have the most structural variation among them, is also encouraging.

Furthermore, while the values of λ we estimate are often quite large, the overall frequencies the model predicts are close to the data. That means that a near-best-response model with a mixture of thinking steps can fit a little better than a QRE model which assumes stochastic response but has only one "type'. The heterogeneity may therefore enable modelers to use best- response calculation and still make probabilistic predictions, which is enormously helpful analytically.

Figures 2 and 3 show how accurately the thinking steps and Nash models fit the data from the three matrix-game data sets. In each Figure, each data point is a separate strategy from each of the games. Figure 2 shows that the data and fits are reasonably good. Figure 3 shows that the Nash predictions (which are often zero or one, pure equilibria, are reasonably accurate though not as close as the thinking-model predictions). Since τ is consistently around 1-2, the thinking model with a single τ could be an adequate approximation to first-period behavior in many different games. To see how far the model can take us, we investigated it in two other classes of games– games with mixed equilibria, and binary entry games. The next section describes results from entry games (see Appendix for details on mixed games).

[13]While the common-τ models have one more free parameter than QRE, any reasonable information criterion penalizing the LL would select the thinking model.

[14]When λ is set to 100 the τ estimates become very regular, around two, which suggests that the variation in estimates is due to poor identification in these games.

[15]The differences in LL across game-specific and common τ are .5, 49.1, 9.4. These are marginally significant (except for Cooper-Van Huyck).

2.2 Market entry games

Consider binary entry games in which there is capacity c (expressed as a fraction of the number of entrants). Each of many entrants decides simultaneously whether to enter or not. If an entrant thinks that fewer than c% will enter she will enter; if she thinks more than c% will enter she stays out.

There are three regularities in many experiments based on entry games like this one (see Ochs, 1999; Seale and Rapoport, 1999; Camerer, 2002, chapter 7): (1) Entry rates across different capacities c are closely correlated with entry rates predicted by (asymmetric) pure equilibria or symmetric mixed equilibria; (2) players slightly over-enter at low capacities and under-enter at high capacities; and (3) many players use noisy cutoff rules in which they stay out for most capacities below some cutoff c^* and enter for most higher capacities.

Let's apply the thinking model with best response. Step zero's enter half the time. This means that when $c < .5$ one step thinkers stay out and when $c > .5$ they enter. Players doing two steps of thinking believe the fraction of zero steppers is $f(0)/(f(0) + f(1)) = 1/(1+\tau)$. Therefore, they enter only if $c > .5$ and $c > \frac{.5+\tau}{1+\tau}$, or when $c < .5$ and $c > \frac{.5}{1+\tau}$. To make this more concrete, suppose $\tau = 2$. Then two-step thinkers enter when $c > 5/6$ and $1/6 < c < 0.5$. What happens is that more steps of thinking "iron out" steps in the function relating c to overall entry. In the example, one-step players are afraid to enter when $c < 1/2$. But when c is not too low (between 1/6 and .5) the two-step thinkers perceive room for entry because they believe the relative proportion of zero-steppers is 1/3 and those players enter half the time. Two-step thinkers stay out for capacities between .5 and 5/6, but they enter for $c > 5/6$ because they know half of the (1/3) zero-step types will randomly stay out, leaving room even though one-step thinkers always enter. Higher steps of thinking smooth out steps in the entry function even further.

The surprising experimental fact is that players can coordinate entry reasonably well, even in the first period. ("To a psychologist," Kahneman (1988) wrote, "this looks like magic".) The thinking steps model provides a possible explanation for this magic and can account for the other two regularities for reasonable τ values. Figure 4 plots entry rates from the first block of two studies for a game similar to the one above (Sundali et al., 1995; Seale and Rapoport, 1999). Note that the number of actual entries rises almost monotonically with c, and entry is above capacity at low c and below capacity at high c.

Figure 4 also shows the thinking steps entry function $N(all|\tau)(c)$ for $\tau = 1.5$ and 2. Both functions reproduce monotonicity and the over- and under- capacity effects. The thinking-steps models also produces approximate cutoff rule behavior for all higher thinking steps except two. When $\tau = 1.5$, step 0 types randomize, step 1 types enter for all c above .5, step 3-4 types use cutoff rules with one "exception", and levels 5-above use strict cutoff rules. This mixture of random, cutoff and near-cutoff rules is roughly what is observed in the data when individual patterns of entry across c are measured (e.g., Seale and Rapoport, 1999).

2.3 Thinking steps and cognitive measures

Since the thinking steps model is a cognitive model, it gives an account of some treatment effects and shows how cognitive measures, like response times and information acquisition, can be correlated with choices.

1. Belief-prompting: Several studies show that asking players for explicit beliefs about what others will do moves their choices, moving them closer to equilibrium (compared to a control in which beliefs are not prompted). A simple example reported in Warglien, Devetag and Legrenzi (1998) is shown in Table 2. Best-responding one-step players think others are randomizing, so they will choose X, which pays 60, rather than Y which has an expected payoff of 45. Higher-step players choose Y.

   Without belief-prompting 70% of the row players choose X. When subjects are prompted to articulate a belief about what the column players will do, 70% choose the dominance-solvable equilibrium choice Y. Croson (2000) reports similar effects. In experiments on beauty contest games, we found that prompting beliefs also reduced dominance-violating choices modestly. Schotter et al. (1994) found a related display effect-showing a game in an extensive-form tree led to more subgame perfect choices.

   Belief-prompting can be interpreted as increasing all players' thinking by one step. To illustrate, assume that since step 0's are forced to articulate *some* belief, they move to step 1. Now they believe others are random so they choose X. Players previously using one or more steps now use two or more. They believe column players choose L so they choose Y. The fraction of X play is therefore due to former

Table 2: How belief-prompting promotes dominance-solvable choices by row players (Warglien, Devetag and Legrenzi, 1998)

	column player		without belief	with belief
row move	L	R	prompting	prompting
X	60,20	60,10	.70	.30
Y	80,20	10,10	.30	.70

zero-step thinkers who now do one step of thinking. This is just one simple example, but the numbers match up reasonably well[16] and it illustrates how belief-prompting effects could be accommodated within the thinking-steps model.

2. Information look-ups: Camerer et al. (1993), Costa-Gomes, Crawford, and Broseta (2001), Johnson et al. (2002), and Salmon (1999) directly measure the information subjects acquire in a game by putting payoff information in boxes which must be clicked open using a computer mouse. The order in which boxes are open, and how long they are open, gives a "subject's-eye view" of what players are looking at, and should be correlated with thinking steps. Indeed, Johnson et al. show that how much time players spend looking ahead to future "pie sizes" in alternating-offer bargaining is correlated with the offers they make. Costa-Gomes et al. show that lookup patterns are correlated with choices that result from various (unobserved) decision rules in normal-form games. These correlations means that a researcher who simply knew what a player had looked at could, to some extent, forecast that player's offer or choice. Both studies also showed that information lookup statistics helped answer questions that choices alone could not.[17]

[16] Take the overconfidence $k-1$ model. The 70% frequency of X choices without belief-prompting is consistent with this model if $f(0|\tau)/2 + f(1|\tau) = .70$, which is most closely satisfied when $\tau = .55$. If belief-prompting moves all thinking up one step, then the former zero-steppers will choose X and all others choose Y. When $\tau = .55$ the fraction of level 0's is 29%, so this simple model predicts 29% choice of X after belief-prompting, close to the 30% that is observed.

[17] Information measures are crucial to resolving the question of whether offers which are close to equal splits are equilibrium offers which reflect fairness concerns, or reflect limited lookahead and heuristic reasoning. The answer is both (see Camerer et al., 1993; Johnson et al., in press. In the Costa-Gomes study, two different decision rules always led to the same choices in their games, but required different lookup patterns. The lookup data were able to therefore classify players according to decision rules more conclusively than choices alone could.

2.4 Summary

A simple model of thinking steps attempts to predict choices in one-shot games and provide initial conditions for learning models. We propose a model which incorporate discrete steps of thinking, and the frequencies of players using different numbers of steps is Poisson-distributed with mean τ. We assume that players at level $K > 0$ cannot imagine players at their level or higher, but they understand the relative proportions of lower-step players and normalize them to compute expected payoffs. Estimates from three experiments on matrix games show reasonable fits for τ around 1-2, and τ is fairly regular across games in two of three data sets. Values of $\tau = 1.5$ also fits data from 15 games with mixed equilibria and reproduces key regularities from binary entry games. The thinking steps model also creates natural heterogeneity across subjects. When best response is assumed, the model generally creates "purification" in which most players at any step level use a pure strategy, but a mixture results because of the mixture of players using different numbers of steps.

3 Learning

By the mid-1990s, it was well-established that simple models of learning could explain some movements in choice over time in specific game and choice contexts.[18] The challenge taken up since then is to see how well a specific parametric model can account for finer details of the equilibration process in wide range of classes of games.

This section describes a one-parameter theory of learning in decisions and games called functional EWA (or fEWA for short; also called "EWA Lite" to emphasize its 'low-calorie' parsimony). fEWA predicts the time path of individual behavior in any normal-form game. Initial conditions can be imposed or estimated in various ways. We use initial conditions from the thinking steps model described in the previous section. The goal is to predict both initial conditions and equilibration in new games in which behavior has never been observed, with minimal free parameters (the model uses two, τ and λ).

[18] To name only a few examples, see Camerer (1987) (partial adjustment models); Smith, Suchanek and Williams (1988) (Walrasian excess demand); McAllister (1991) (reinforcement); Camerer and Weigelt (1993) (entrepreneurial stockpiling);Roth and Erev (1995) (reinforcement learning); Ho and Weigelt (1996) (reinforcement and belief learning); Camerer and Cachon (1996) (Cournot dynamics).

3.1 Parametric EWA learning: Interpretation, uses and limits

fEWA is a relative of a parametric model of learning called experience-weighted attraction (EWA) (Camerer and Ho 1998, 1999). As in most theories, learning in EWA is characterized by changes in (unobserved) attractions based on experience. Attractions determine the probabilities of choosing different strategies through a logistic response function. For player i, there are m_i strategies (indexed by j) which have initial attractions denoted $A_i^j(0)$. The thinking steps model is used to generate initial attractions given parameter values τ and λ.

Denote i's j'th strategy by s_i^j, chosen strategies by i and other players (denoted $-i$) by $s_i(t)$ and $s_{-i}(t)$, and player i's payoffs by $\pi_i(s_i^j, s_{-i}(t))$.[19] Define an indicator function $I(x, y)$ to be zero if $x \neq y$ and one if $x = y$. The EWA attraction updating equation is

$$A_i^j(t) = \frac{\phi N(t-1) A_i^j(t-1) + [\delta + (1-\delta) I(s_i^j, s_i(t))] \pi_i(s_i^j, s_{-i}(t))}{N(t-1)\phi(1-\kappa) + 1} \tag{3.1}$$

and the experience weight (the "EW" part) is updated according to $N(t) = N(t-1)\phi(1-\kappa) + 1$.

Notice that the term $[\delta + (1-\delta) I(s_i^j, s_i(t))]$ implies that a weight of one is put on the payoff term when the strategy being reinforced is the one the player chose ($s_i^j = s_i(t)$), but the weight on foregone payoffs from unchosen strategies ($s_i^j \neq s_i(t)$) is δ. Attractions are mapped into choice probabilities using a logit response function $P_i^j(t+1) = \frac{e^{\lambda \cdot A_i^j(t)}}{\sum_{k=1}^{m_i} e^{\lambda \cdot A_i^k(t)}}$ (where λ is the response sensitivity). The subscript i, superscript j, and argument $t+1$ in $P_i^j(t+1)$ are reminders that the model aims to explain every choice by every subject in every period.[20]

Each EWA parameter has a natural interpretation.

The parameter δ is the weight placed on foregone payoffs. It presumably is affected by imagination (in psychological terms, the strength of counterfactual reasoning or regret, or in economic terms, the weight placed on opportunity costs and benefits) or reliability of information about foregone payoffs (Heller and Sarin, 2000).

[19] To avoid complications with negative payoffs, we rescale payoffs by subtracting by the minimum payoff so that rescale payoffs are always weakly positive.

[20] Other models aim to explain choices aggregated at some level. Of course, models of this sort can sometimes be useful. But our view is that a parsimonious model which can explain very fine-grained data can probably explain aggregated data well too, but the opposite may not be true.

The parameter ϕ decays previous attractions due to forgetting or, more interestingly, because agents are aware that the learning environment is changing and deliberately "retire" old information (much as firms junk old equipment more quickly when technology changes rapidly).

The parameter κ controls the rate at which attractions grow. When $\kappa = 0$ attractions are weighted averages and grow slowly; when $\kappa = 1$ attractions cumulate. We originally included this variable because some learning rules used cumulation and others used averaging. It is also a rough way to capture the distinction in machine learning between "exploring" an environment (low κ), and "exploiting" what is known by locking in to a good strategy (high κ) (e.g., Sutton and Barto, 1998).

The initial experience weight $N(0)$ is like a strength of prior beliefs in models of Bayesian belief learning. It plays a minimal empirical role so it is set to one in our current work.

EWA is a hybrid of two widely-studied models, reinforcement and belief learning. In reinforcement learning, only payoffs from chosen strategies are used to update attractions and guide learning. In belief learning, players do not learn about which strategies work best; they learn about what others are likely to do, then use those updated beliefs to change their attractions and hence what strategies they choose (see Brown, 1951; Fudenberg and Levine, 1998). EWA shows that reinforcement and belief learning, which were often treated as fundamentally different, are actually related in a non-obvious way, because both are special kinds of reinforcement rules.[21] When $\delta = 0$ the EWA rule is a simple reinforcement rule[22]. When $\delta = 1$ and $\kappa = 0$ the EWA rule is equivalent to belief learning using weighted fictitious play.[23]

Foregone payoffs are the fuel that runs EWA learning. They also provide an indirect link to "direction learning" and imitation. In direction learning players move in the direction of observed best response (Selten and Stöcker, 1986). Suppose players follow EWA

[21] See also Cheung and Friedman, 1997, pp. 54-55; Fudenberg and Levine, 1998, pp. 184-185; and Ed Hopkins, in press.

[22] See Bush and Mosteller, 1955; Harley, 1981; Cross, 1983; Arthur, 1991; McAllister, 1991; Roth and Erev, 1995; Erev and Roth, 1998.

[23] When updated fictitious play beliefs are used to update the expected payoffs of strategies, precisely the same updating is achieved by reinforcing all strategies by their payoffs (whether received or foregone). The belief themselves are an epiphenomenon that disappear when the updating equation is written expected payoffs rather than beliefs.

but don't know foregone payoffs, and believe those payoffs are monotonically increasing between their choice $s_i(t)$ and the best response. If they also reinforce strategies near their choice $s_i(t)$ more strongly than strategies that are further away, their behavior will look like direction learning. Imitating a player who is similar and successful can also be seen as a way of heuristically inferring high foregone payoffs from an observed choice and moving in the direction of those higher payoffs.

The relation of various learning rules can be shown visually in a cube showing configurations of parameter values (see Figure 5). Each point in the cube is a triple of EWA parameter values which specifies a precise updating equation. The corner of the cube with $\phi = \kappa = 0, \delta = 1$, is Cournot best-response dynamics. The corner $\kappa = 0, \phi = \delta = 1$, is standard fictitious play. The vertex The relation of various learning rules can be shown visually in a cube showing configurations of parameter values (see Figure 5). Each point in the cube is a triple of EWA parameter values which specifies a precise updating equation. The corner of the cube with $\phi = \kappa = 0, \delta = 1$, is Cournot best-response dynamics. The corner $\kappa = 0, \phi = \delta = 1$, is standard fictitious play. The vertex connecting these corners, $\delta = 1, \kappa = 0$, is the class of weighted fictitious play rules (e.g., Fudenberg and Levine, 1998). The vertices with $\delta = 0$ and $\kappa = 0$ or 1 are averaging and cumulative choice reinforcement rules (Roth and Erev, 1995; and Erev and Roth, 1998).

The biologist Francis Crick (1988) said, "in nature a hybrid is often sterile, but in science the opposite is usually true". As Crick suggests, the point of EWA is not simply to show a surprising relation among other models, but to improve their fertility for explaining patterns in data by combining the best modeling "genes". In reinforcement theories received payoffs get the most weight (in fact, all the weight[24]). Belief theories implicitly assume that foregone and received payoffs are weighted equally. Rather than assuming one of these intuitions about payoff weights is right and the other is wrong, EWA allows *both* intuitions to be true. When $0 < \delta < 1$ received payoffs can get *more* weight, but foregone payoffs also get *some* weight.

The EWA model has been estimated by ourselves and many others on about 40 data sets (see Camerer, Hsia, and Ho, 2000). The hybrid EWA model predicts more accurately than the special cases of reinforcement and weighted fictitious in most cases, except in

[24]Taken seriously, reinforcement models also predict that learning paths will look the same whether players know their full payoff matrix or not. This prediction is rejected in all the studies that have tested it, e.g., Mookerjhee and Sopher, 1994; Rapoport and Erev, 1998; Battalio, Van Huyck, and Rankin, 2001.

games with mixed-strategy equilibrium where reinforcement does equally well.[25] In our model estimation and validation, we always penalize the EWA model in ways that are known to make the *adjusted* fit *worse* if a model is too complex (i.e., if the data are actually generated by a simpler model).[26] Furthermore, econometric studies show that if the data were generated by simpler belief or reinforcement models, then EWA estimates would correctly identify that fact for most games and reasonable sample sizes (see Salmon, 2001; Cabrales and Garcia-Fontes, 2000). Since EWA is capable of identifying behavior consistent with special cases, when it does not then the hybrid parameter values are improving fit.

Figure 5 also shows estimated parameter triples from twenty data sets. Each point is an estimate from a different game. If one of the special case theories is a good approximation to how people generally behave across games, estimated parameters should cluster in the corner or vertex corresponding to that theory. In fact, parameters tend to be sprinkled around the cube, although many (typically mixed-equilibrium games) cluster in the averaged reinforcement corner with low δ and κ. The dispersion of estimates in the cube raises an important question: Is there regularity in which games generate which parameter estimates? A positive answer to this question is crucial for predicting behavior in brand new games.

This concern is addressed by a version of EWA, fEWA, which replaces free parameters with deterministic functions $\phi_i(t), \delta_i(t), \kappa_i(t)$ of player i's experience up to period t. These functions determine parameter values for each player and period. The parameter values are then used in the EWA updating equation to determine attractions, when then determine choices probabilistically. Since the functions also vary across subjects and over time, they have the potential to inject heterogeneity and time-varying "rule learning", and to explain learning better than models with fixed parameter values across people and time. And since fEWA has only one parameter which must be estimated (λ)[27], it is especially helpful when learning models are used as building blocks for more complex

[25] In mixed games no model improves much on Nash equilibrium (and often don't improve on quantal response equilibrium at all), and parameter identification is poor; see Salmon, 2001))

[26] We typically penalize in-sample likelihood functions using the Akaike and Bayesian information criteria, which subtract a penalty of one, or log(n), times the number of degrees of freedom from the maximized likelihood. More persuasively, we rely mostly on out-of-sample forecasts which will be *less* accurate if a more complex model simply appears to fit better because it overfits in-sample.

[27] Note that if your statistical objective is to maximize hit rate, λ does not matter and so fEWA is a zero-parameter theory given initial conditions.

models that incorporate sophistication (some players think others learn) and teaching, as we discuss in the section below.

The crucial function in fEWA is $\phi_i(t)$, which is designed to detect change in the learning environment. As in physical change detectors, such as security systems or smoke alarms, the challenge is to detect change when it is really occurring, but not falsely mistake noise for change too often. The core of the function is a "surprise index", the difference between the other players' strategies in the window of the last W periods and the average strategy of others in all previous periods (where W is the minimal support of Nash equilibria, smoothing fluctuations in mixed games). The function is specified in terms of relative frequencies of strategies, without using information about how strategies are ordered, but is easily extended to ordered strategies (like prices or locations). Change is measured by taking the differences in corresponding elements of the two frequency vectors (recent history and all history), squaring them, and sum over strategies. Dividing by two and subtracting from one normalizes the function so it is between zero and one and is smaller when change is large. The change-detection function $\phi_i(t)$ is

$$\phi_i(t) = 1 - .5(\sum_{j=1}^{m_{-i}}[\frac{\sum_{\tau=t-W+1}^{t} I(s^j_{-i}, s_{-i}(\tau))}{W} - \frac{\sum_{\tau=1}^{t} I(s^j_{-i}, s_{-i})}{t}]^2) \tag{3.2}$$

The term $\frac{\sum_{\tau=t-W+1}^{t} I(s^j_{-i}, s_{-i}(\tau))}{W}$ is the j-th element of a vector that simply counts how often strategy j was played by the others in periods $t-W+1$ to t, and divides by W. The term $\frac{\sum_{\tau=1}^{t} I(s^j_{-i}, s_{-i})}{t}$ is the relative frequency count of the j-th strategy over all t periods.[28] When recent observations of what others have done deviate a lot from all previous observations, the deviations in strategy frequencies will be high and ϕ will be low. When recent observations are like old observations, ϕ will be high. Since a very low ϕ erases old history– permanently– ϕ should be kept close to one unless there is an unmistakable change in what others are doing. The function above only dips toward zero if a single strategy has been played by others in all $t-1$ previous periods and then a new strategy is played. (Then $\phi_i(t) = \frac{2t-1}{t^2}$, which is .75, .56 and .19 for t=2,3,10.)[29]

[28] In games with multiple players, the frequency count of the relevant aggregate statistics is used. For example, in median action game, the frequency count of the median strategy by all other players in each period is used.

[29] Another interesting special case is when different strategies have been played in every period up to $t-1$, and another different strategy is played. (This is often true in games with large strategy spaces, such as location or pricing, when order of strategies is not used.) Then $\phi_i(t) = .5 + \frac{1}{2t}$, which starts at .75 and asymptotes at .5.

The other fEWA functions are less empirically important and interesting so we mention them only briefly. The function $\delta_i(t) = \phi_i(t)/W$. Dividing by W pushes $\delta_i(t)$ toward zero in games with mixed equilibria, which matches estimates in many games (see Camerer, Ho and Chong, in press).[30] Tying $\delta_i(t)$ to the change detector $\phi_i(t)$ means chosen strategies are reinforced relatively strongly (compared to unchosen ones) when change is fast. This reflects a "status quo bias" or "freezing" response to danger (which is virtually universal across species, including humans). Since $\kappa_i(t)$ controls how sharply subjects lock in to choosing a small number of strategies, we use a "Gini coefficient"– a standard measure of dispersion often used to measure income inequality– over choice frequencies [31]

fEWA has three advantages. First, it is easy to use because it has only one free parameter (λ). Second, parameters in fEWA naturally vary across time and people (as well as across games), which can capture heterogeneity and mimic "rule learning" in which parameters vary over time (e.g., Stahl, 1996, 2000, and Salmon, 1999). For example, if ϕ rises across periods from 0 to 1 as other players stabilize, players are effectively switching from Cournot-type dynamics to fictitious play. If δ rises from 0 to 1, players are effectively switching from reinforcement to belief learning. Third, it should be easier to theorize about the limiting behavior of fEWA than about some parametric models. A key feature of fEWA is that as a player's opponents' behavior stabilizes, $\phi_i(t)$ goes toward one and (in games with pure equilibria) $\delta_i(t)$ does too. If $\kappa = 0$, fEWA then automatically turns into fictitious play; and a lot is known about theoretical properties of fictitious play.

3.2 fEWA predictions

In this section we compare in-sample fit and out-of-sample predictive accuracy of different learning models when parameters are freely estimated, and check whether fEWA functions can produce game-specific parameters similar to estimated values. We use seven games: Games with unique mixed strategy equilibrium (Mookerjhee and Sopher, 1997); R&D patent race games (Rapoport and Amaldoss, 2000); a median-action order statistic coordination game with several players (Van Huyck, Battalio, and Beil, 1991); a continental-divide coordination game, in which convergence behavior is extremely sen-

[30] If one is uncomfortable assuming subjects act as if they know W, you can easily replace W by some function of the variability of others' choices to proxy for W.

[31] Formally, $\kappa_i(t) = 1 - 2 \cdot \{\sum_{k=1}^{m_i} f_i^{(k)}(t) \cdot \frac{m_i - k}{m_i - 1}\}$ where $f_i^k(t)$ are ranked from the lowest to the highest.

sitive to initial conditions (Van Huyck, Cook, and Battalio, 1997); a "pots game" with entry into two markets of different sizes (Amaldoss and Ho, in preparation); dominance-solvable p-beauty contests (Ho, Camerer, and Weigelt, 1998); and a price-matching game (called "travellers' dilemma" by Capra, Goeree, Gomez and Holt, 2000).

3.3 Estimation Method

The estimation procedure for fEWA is sketched briefly here (see Ho, Camerer, and Chong, 2001 for details). Consider a game where N subjects play T rounds. For a given player i of level c, the likelihood function of observing a choice history of $\{s_i(1), s_i(2), \ldots, s_i(T-1), s_i(T)\}$ is given by:

$$\Pi_{t=1}^{T} P_i^{s_i(t)}(t|c) \tag{3.3}$$

The joint likelihood function L of observing all players' choice is given by

$$L(\lambda) = \Pi_i^N \{\sum_{c=1}^{K} f(c) \cdot \Pi_{t=1}^{T} P_i^{s_i(t)}(t)\} \tag{3.4}$$

where K is set to a multiple of τ rounded to an integer. Most models are "burned in" by using first-period data to determine initial attractions. We also compare all models with burned-in attractions with a model in which the thinking steps model from the previous section is used to create initial conditions and combined with fEWA. Note that the latter hybrid uses only two parameters (τ and λ) and does not use first-period data at all.

Given the initial attractions and initial parameter values[32], attractions are updated using the EWA formula. fEWA parameters are then updated according to the functions above and used in the EWA updating equation. Maximum likelihood estimation is used to find the best-fitting value of λ (and other parameters, for the other models) using data from the first 70% of the subjects. Then the value of λ is frozen and used to forecast behavior of the entire path of the remaining 30% of the subjects. Payoffs were all converted to dollars (which is important for cross-game forecasting).

In addition to fEWA (one parameter), we estimated the parametric EWA model (five parameters), a belief-based model (weighted fictitious play, two parameters) and the two

[32]The initial parameter values are $\phi_i(0) = \kappa_i(0) = .5$ and $\delta_i(0) = \phi_i(0)/W$. These initial values are averaged with period-specific values determined by the functions, weighting the initial value by $\frac{1}{t}$ and the functional value by $\frac{t-1}{t}$.

Table 3: Out of sample accuracy of learning models (Ho, Camerer and Chong, 2001)

	Thinking +fEWA		fEWA		EWA		Weighted fict. play		Reinf. with PV		QRE	
game	%Hit	LL	%Hit	LL	%Hit	LL	%Hit	LL	%Hit	LL	%Hit	LL
Cont'l divide	**45**	-483	**47**	-470	**47**	**-460**	25	-565	45	-557	5	-806
Median action	71	-112	74	-104	**79**	**-83**	**82**	-95	74	-105	49	-285
p-BC	**8**	-2119	**8**	-2119	**6**	**-2042**	**7**	-2051	**6**	-2504	4	-2497
Price matching	**43**	-507	**46**	-445	**43**	**-443**	36	-465	**41**	-561	27	-720
Mixed games	**36**	-1391	**36**	**-1382**	**36**	-1387	**34**	-1405	33	-1392	**35**	-1400
Patent Race	**64**	-1936	**65**	-1897	**65**	-1878	53	-2279	**65**	**-1864**	40	-2914
Pot Games	**70**	-438	**70**	-436	**70**	-437	66	-471	**70**	**-429**	51	-509
Pooled	**50**	-6986	**51**	**-6852**	49	-7100	40	-7935	46	-9128	36	-9037
KS p-BC			6	-309	3	-279	3	-279	4	-344	1	-346

Note: Sample sizes are 315, 160, 580, 160, 960, 1760, 739, 4674 (pooled), 80.

-parameter reinforcement models with payoff variability (Erev, Bereby-Meyer and Roth, 1999; Roth et al., 2000), and QRE.

3.4 Model fit and predictive accuracy in all games

The first question we ask is how well models fit and predict on a game-by-game basis (i.e., parameters are estimated separately for each game). For out-of-sample validation we report both hit rates (the fraction of most-likely choices which are picked) and log likelihood (LL). (Keep in mind that these results forecast a holdout sample of subjects after model parameters have been estimated on an earlier sample and then "frozen". If a complex model is fitting better within a sample purely because of spurious overfitting, it will predict more poorly out of sample.) Results are summarized in Table 3.

The best fits for each game and criterion are printed in bold; hit rates which statistically indistinguishable from the best (by the McNemar test) are also in bold. Across games, parametric EWA is as good as all other theories or better, judged by hit rate, and has the best LL in four games. fEWA also does well on hit rate in six of seven games. Reinforcement is competitive on hit rate in five games and best in LL in two. Belief models are often inferior on hit rate and never best in LL. QRE clearly fits worst.

Combining fEWA with a thinking steps model to predict initial conditions (rather than using the first-period data), a two-parameter combination, is only a little worse in hit rate than fEWA and slightly worse in LL.

The bottom line of Table 3, "pooled", shows results when a single set of common parameters is estimated for all games (except for game-specific λ). If fEWA is capturing parameter differences across games effectively, it should predict especially accurately, compared to other models, when games are pooled. It does: When all games are pooled, fEWA predicts out-of-sample better than other theories, by both statistical criteria.

Some readers of our functional EWA paper were concerned that by searching across different specifications, we may have overfitted the sample of seven games we reported. To check whether we did, we announced at conferences in 2001 that we would analyze all the data people sent us by the end of the year and report the results in a revised paper. Three samples were sent and we analyzed one so far– experiments by Kocher and Sutter (2000) on p-beauty contest games played by individuals and groups. The KS results are reported in the bottom row of Table 3. The game is the same as the beauty contests we studied (except for the interesting complication of group decision making, which speeds equilibration), so it is not surprising that the results replicate the earlier findings: Belief and parametric EWA fit best by LL, followed by fEWA, and reinforcement and QRE models fit worst. This is a small piece of evidence that the solid performance of fEWA (while worse than belief learning on these games) is not entirely due to overfitting on our original 7-game sample.

The Table also shows results (in the column headed "Thinking+fEWA") when the initial conditions are created by the thinking steps model rather than from first-period data and combined with the fEWA learning model. Thinking plus fEWA are also a little more accurate than the belief and reinforcement models in five of seven games. The hit rate and LL suffer only a little compared to the fEWA with estimated parameters. When common parameters are estimated across games (the row labelled "pooled"), fixing initial conditions with the thinking steps model only lowers fit slightly.

Now we will show predicted and relative frequencies for three games which highlight differences among models. In other games the differences are minor or hard to see with the naked eye.[33]

[33]More details are in Ho, Camerer and Chong, 2001, and corresponding graphs for *all* games can be seen at http://www.fba.nus.edu.sg/depart/mk/fbacjk/ewalite/ewalite.htm

3.5 Dominance-solvable games: Beauty contests

In beauty contest games each of n players chooses $x_i \in [0, 100]$. The average of their choices is computed and whichever player is closest to $p < 1$ times the average wins a fixed prize (see Nagel, 1999, for a review). The unique Nash equilibrium is zero. (The games get their name from a passage in Keynes about how the stock market is like a special beauty contest in which people judge who others will think is beautiful.) These games are a useful way to measure the steps of iterated thinking players seem to use (since higher steps will lead to lower number choices). Experiments have been run with exotic subject pools like Ph.D's and CEOs (Camerer, 1997), and in newspaper contests with very large samples (Nagel et al., 1999). The results are generally robust although specially-educated subjects (e.g., professional game theorists) choose, not surprisingly, closer to equilibrium.

We analyze experiments run by Ho, Camerer and Weigelt (1998).[34] The data and relative frequencies predicted by each learning model are shown in Figure 6a-f. Figure 6a shows that while subjects start around the middle of the distribution, they converge downward steadily toward zero. By period 5 half the subjects choose numbers 1-10.

The EWA, belief, and thinking-fEWA model all capture the basic regularities although they underestimate the speed of convergence. (In the next section we add sophistication– some subjects know that others are learning and "shoot ahead" of the learners by choosing lower numbers– which improves the fit substantially.) The QRE model is a dud in this game and reinforcement also learns far too slowly because most players receive no reinforcement.[35]

[34]Subjects were 196 undergraduate students in computer science and engineering in Singapore. Each group played 10 times together twice, with different values of p in the two 10-period sequences. (One sequence used $p > 1$ and is not included.) We analyze a subsample of their data with $p = .7$ and .9, from groups of size 7. This subsample combines groups in a 'low experience' condition (the game is the first of two they play) and a 'high experience' condition (the game is the second of two, following a game with $p > 1$).

[35]Reinforcement can be sped up in such games by reinforcing unchosen strategies in some way, e.g., Roth and Erev, 1995, which is why EWA and belief learning do better.

Table 4: Payoffs in 'continental divide' experiment, Van Huyck, Cook and Battalio (1997)

	Median Choice													
choice	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	45	49	52	55	56	55	46	-59	-88	-105	-117	-127	-135	-142
2	**48**	53	58	62	65	66	61	-27	-52	-67	-77	-86	-92	-98
3	**48**	**54**	***60***	**66**	70	74	72	1	-20	-32	-41	-48	-53	-58
4	43	51	58	65	**71**	**77**	80	26	8	-2	-9	-14	-19	-22
5	35	44	52	60	69	**77**	**83**	46	32	25	19	15	12	10
6	23	33	42	52	62	72	82	62	53	47	43	41	39	38
7	7	18	28	40	51	64	78	75	69	66	64	63	62	62
8	-13	-1	11	23	37	51	69	83	81	80	80	80	81	82
9	-37	-24	-11	3	18	35	57	88	89	91	92	94	96	98
10	-65	-51	-37	-21	-4	15	40	**89**	**94**	98	101	104	107	110
11	-97	-82	-66	-49	-31	-9	20	85	**94**	**100**	105	110	114	119
12	-133	-117	-100	-82	-61	-37	-5	78	91	99	**106**	***112***	**118**	**123**
13	-173	-156	-137	-118	-96	-69	-33	67	83	94	103	110	117	**123**
14	-217	-198	-179	-158	-134	-105	-65	52	72	85	95	104	112	120

3.6 Games with multiple equilibria: Continental divide game

Van Huyck, Cook and Battalio (1997) studied a coordination game with multiple equilibria and extreme sensitivity to initial conditions, which we call the continental divide game (CDG). The payoffs in the game are shown in Table 4. Subjects play in cohorts of seven people. Subjects choose an integer from 1 to 14, and their payoff depends on their own choice and on the median choice of all seven players.

The payoff matrix is constructed so that there are two pure equilibria (at 3 and 12) which are Pareto-ranked (12 is the better one). Best responses to different medians are in bold. The best-response correspondence bifurcates in the middle: If the median starts at 7 virtually any sort of learning dynamics will lead players toward the equilibrium at 3. If the median starts at 8 or above, however, learning will eventually converge to an equilibrium of 12. Both equilibrium payoffs are shown in bold italics. The payoff at 3 is about half as much as at 12 so which equilibrium is selected has a large economic impact.

Figures 7a-f show empirical frequencies (pooling all subjects) and model predictions.[36] The key features of the data are: Bifurcation over time from choices in the middle of the range (5-10) to the extremes, near the equilibria at 3 and 12; and late-period choices are more clustered around 12 than around 3. There is also an extreme sensitivity to initial conditions (which is disguised by the aggregation across sessions in Figure 7a): Namely, five groups had initial medians below 7 and all five converged toward the inefficient low equilibrium. The other five groups had initial medians above 7 and all five converged toward the efficient high equilibrium. This path-dependence shows the importance of a good theory of initial conditions (such as the thinking steps model). Because a couple of steps of thinking generates a distribution concentrated in the middle strategies 5-9, the thinking-steps models predicts that initial medians will sometimes be above the separatrix 7 and sometimes below. The model does not predict precisely *which* equilibrium will emerge, but it predicts that both high and low equilibria will sometimes emerge.

Notice also that strategies 1-4 are never chosen in early periods, but are frequently chosen in later periods. Strategies 7-9 are frequently chosen in early periods but rarely chosen in later periods. Like a sportscar, a good model should be able to capture these effects by "accelerating" low choices quickly (going from zero to frequent choices in a few periods) and "braking" midrange choices quickly (going from frequent choices to zero).

QRE fits poorly because it predicts no movement (it is not a theory of learning, of course, but simply a static benchmark which is tougher to beat than Nash). Reinforcement with PV fits well. Belief learning does not reproduce the asymmetry between sharp convergence to the high equilibrium and flatter frequencies around the low equilibrium. The reason why is diagnostic of a subtle weakness in belief learning. Note from Table 4 that the payoff gradients around the equilibria at 3 and 12 are exactly the same– choosing one number too high or low "costs" \$.02; choosing two numbers too high or low costs \$.08, and so forth. Since belief learning computes expected payoffs, and the logit rule means only differences in expected payoffs influence choice probability, the fact that the payoff gradients are the same means the spread of probability around the two equilibria must be the same. fEWA, parametric EWA, and the reinforcement models generate the asymmetry with low δ.[37]

[36] Their experiment used 10 cohorts of seven subjects each, playing for 15 periods. At the end of each period subjects learned the median, and played again with the same group in a partner protocol. Payoffs were the amounts in the table, in pennies.

[37] At the high equilibrium, the payoffs are larger and so the difference between the received payoff and δ times the foregone payoff will be larger than at the low equilibrium. (Numerically, a player who chooses

3.7 Games with dominance-solvable equilibrium: Price-matching with loyalty

Capra et al. (1999) studied a dominance-solvable price-matching game. In their game two players simultaneously choose a price between 80 and 200. Both players earn the low price. In addition, the player who names the lower price receives a bonus of R and the players who names the higher price pays a penalty R. (If their prices are the same the bonus and penalty cancel and players just earn the price they named.) You can think of R as a reduced-form expression of the benefits of customer loyalty and word-of-mouth which accrue to the lower-priced player, and the penalty is the cost of customer disloyalty and switching away from the high-price firm. We like this game because price-matching is a central feature of economic life. These experiments can also, in principle, be tied to field observations in future work.

Their experiment used six groups of 9-12 subjects. The reward/penalty R had six values (5, 10, 20, 25, 50, 80). Subjects were rematched randomly.[38]

Figures 8a-f show empirical frequencies and model fits for R=50 (where the models differ most). A wide range of prices are named in the first round. Prices gradually fall, between 91-100 in rounds 3-5, 81-90 in rounds 5-6, and toward the equilibrium of 80 in later rounds.

QRE predicts a spike at the Nash equilibrium of 80.[39] The belief-based model predicts the direction of convergence, but overpredicts numbers in the interval 81-90 and underpredicts choices of precisely 80. The problem is that the incentive in the travellers'

3 when the median is 3 earns \$.60 and has a foregone payoff from 2 or 4 of \$.58 $\cdot\delta$. The corresponding figures for a player choosing 12 are \$1.12 and \$1.10 $\cdot\ \delta$. The differences in received and foregone payoffs around 12 and around 3 are the same when $\delta = 1$ but the difference around 12 grows larger as δ falls (for example, for the fEWA estimate $\hat{\delta} = .69$, the differences are \$.20 and \$.36 for 3 and 12.) Cumulating payoffs rather than averaging them "blows up" the difference and produces sharper convergence at the high equilibrium.

[38] They also had a session with $R = 10$ but in this session one subject sat out each round so we dropped it to avoid making an ad hoc assumption about learning in this unusual design. Each subject played 10 times (and played with a different R for five more rounds; we use only the first 10 rounds)

[39] As λ rises, the QRE equilibria move sharply from smearing probability throughout the price range (for low λ) to a sharp spike at the equilibrium (higher λ). No intermediate λ can explain the combination of initial dispersion and sharp convergence at the end so the best-fitting QRE model essentially makes the Nash prediction.

dilemma is to undercut the other player's price by as little as possible. Players only choose 80 frequently in the last couple of periods; before those periods it pays to choose higher numbers.

EWA models explain the sharp convergence in late periods by cumulating payoffs and estimating $\delta = .63$ (for fEWA). Players who chose 80 while others named a higher price could have earned more by undercutting the other price, but weighting that higher foregone payoff by δ means their choice of 80 is reinforced more strongly, which matches the data.

Reinforcement with payoff-variability has a good hit rate because the highest spikes in the graph often correspond with spikes in the data. But the graph shows that predicted learning is much more sluggish than in the data (i.e., the spikes are not high enough). Because $\phi = 1$ and players are not predicted to move toward ex-post best responses, the model cannot explain why players learn to choose 80 so rapidly.

3.8 Economic value of learning models

In the last couple of decades the concept of economic engineering has gradually emerged from its start in the late 1970s (see Plott, 1986) as increasingly important. Experimentation has played an important role in this emergence (see Plott, 1997; Rassenti, Smith and Wilson, 2001; Roth, 2001). For the practice of economic engineering, it is useful to have a measure of how much value a theory or design creates. For policy purposes increases in allocative efficiency are a sensible measure. But for judging the private value of advice to a firm or consumer other measures are more appropriate.

Camerer and Ho (2001) introduced a measure called "economic value". The economic value of a learning theory is how much model forecasts of behavior of other players improve the profitability of a particular player's choices. This measure treats a theory as being like the advice service professionals sell (e.g., consultants). The value of a theory is the difference in the economic value of the client's decisions with and without the advice.

In equilibrium, the economic value of a learning theory is zero by definition. A bad theory, which implicitly "knows" less than the subjects themselves do about what other subjects are likely to do, will have negative economic value.

To measure economic value, we use model parameters and a player's observed experience through period t to generate model predictions about what others will do in t+1. Those predictions are used to compute expected payoffs from strategies and recommend a choice with the highest expected value. We then compare the profit from making that choice in t+1 (given what other players did in t+1) with profit from the target player's actual choice. Economic value is a good measure because it uses the full distribution of predictions about what other players are likely to do, *and* the economic impact of those possible choices. We have not yet controlled for the boomerang effect of how a recommended choice would have changed future behavior by others, but this effect will be small in most of the games.[40]

Data from six games are used to estimate model parameters and make recommendations in the seventh game, for each of the games separately. Table 5 shows the overall economic value– the percentage improvement (or decline) in payoffs of subjects from following a model recommendation rather than their actual choices. The highest economic value for each game is printed in bold. Most models have positive economic value.[41] The percentage improvement is small in some games because even clairvoyant advice would not raise profits much.[42]

fEWA and EWA usually add the most value (except in pot games, where only QRE adds value). Belief learning has positive economic value in all but one game. Reinforcement learning adds the most value in patent races, but has negative economic value in three other games. (Reinforcement underestimates the rate of strategy change in continental divide and beauty contest games, and hence gives bad advice.) QRE has negative

[40]In beauty contests and coordination games, payoffs depend on the mean or median of fairly large groups (7-9 except in 3-person entry games) so switching one subject's choice to the recommendation would probably not change the mean or median and hence would not change future behavior much. In other games players are usually paired randomly so the boomerang effect again is muted. We are currently redoing the analysis to simply compare profits of players whose choices frequently matched the recommendation with those who rarely did. This controls for the boomerang effect and also for a Lucas critique effect in which adopting recommendations would change behavior of others and hence the model parameters used to derive the recommendations. A more interesting correction is to run experiments in which one or more computerized subjects actually use a learning model to make choices, and compare their performance with that of actual subjects.

[41]We are currently working on computing economic value of the thinking plus fEWA specification.

[42]For example, in the continental divide game, ex post optimal payoffs would have been 892 (pennies per player) if players knew exactly what the median would be, and subjects actually earned 837. EWA and fEWA generate simulated profits of 879-882, which is only an improvement of 5% over 837 but is 80% of the maximum possible improvement from actual payoffs to clairvoyant payoffs.

Table 5: Economic value of learning theories (% improvement in payoffs)

Game	functional EWA	parametric EWA	Belief-based	Reinf.-PV	QRE
continental divide	5.0%	**5.2%**	4.6%	-9.4%	-30.4%
median action	1.5%	**1.5%**	1.2%	1.3%	-1.0%
p-Beauty contest	**49.9%**	40.8%	26.7%	-7.2%	-63.5%
price matching	**10.3%**	9.8%	9.4%	3.4%	2.7%
mixed strategies	**7.5%**	3.0%	1.1%	5.8%	-1.8%
patent race	1.7%	1.2%	1.3%	**2.9%**	1.2%
pot games	-2.7%	-1.1%	-1.3%	-1.9%	**9.9%**

economic value in four games.

3.9 Summary

This section reports a comparison among several learning models on seven data sets. The new model is fEWA, a variant of the hybrid EWA model in which estimated parameters are replaced by functions which are entirely determined by data. fEWA captures predictable cross-game variation in parameters and hence fits better than other models when common parameters are estimated across games. A closer look at the continental divide and price-matching games shows that belief models are close to the data on average but miss other features (the asymmetry in convergence toward each of the two pure equilibria in the continental divide game, and the sharp convergence on the minimum price in price-matching). Reinforcement predicts well in coordination games and predicts the correct price often in price-matching (but with too little probability). However, reinforcement predicts badly in beauty contest games. It is certainly true that for explaining some features of some games, the reinforcement and belief models are adequate. But fEWA is easier to estimate (it has one parameter instead of two) and explains subtler features other models sometimes miss. It is also never fits poorly (relative to other games), which is the definition of robustness.

4 Sophistication and teaching

The learning models discussed in the last section are adaptive and backward-looking: Players only respond to their own previous payoffs and knowledge about what others did. While a reasonable approximation, these models leave out two key features: Adaptive players do not explicitly use information about *other* players' payoffs (though subjects actually do[43]); and adaptive models ignore the fact that when the same players are matched together repeatedly, their behavior is often different than then they are not rematched together, generally in the direction of greater efficiency (e.g., Andreoni and Miller (1993), Clark and Sefton (1999), Van Huyck, Battalio and Beil (1990)).

In this section adaptive models are extended to include sophistication and strategic teaching in repeated games (see Stahl, 1999; and Camerer, Ho and Chong, in press, for details). Sophisticated players believe that others are learning and anticipate how others will change in deciding what to do. In learning to shoot a moving target, for example, soldiers and fighter pilots learn to shoot *ahead*, toward where the target *will be*, rather than shoot at the current target. They become sophisticated.

Sophisticated players who also have strategic foresight will "teach"– that is, they choose current actions which teach the learning players what to do, in a way that benefits the teacher in the long-run. Teaching can be either mutually-beneficial (trust-building in repeated games) or privately-beneficial but socially costly (entry-deterrence in chain-store games). Note that sophisticated players will use information about payoffs of others (to forecast what others will do) and will behave differently depending on how players are matched, so adding sophistication can conceivably account for effects of information and matching that adaptive models miss.[44]

4.1 Sophistication

Let's begin with myopic sophistication (no teaching). The model assumes a population mixture in which a fraction α of players are sophisticated. To allow for possible overconfidence, sophisticated players think that a fraction $(1-\alpha')$ of players are adaptive and the

[43] Partow and Schotter (1993), Mookerjee and Sopher (1994), Cachon and Camerer (1996).

[44] Sophistication may also potentially explain why players sometimes move in the *opposite* direction predicted by adaptive models (Rapoport, Lo and Zwick, 1999), and why measured beliefs do not match up well with those predicted by adaptive belief learning models (Nyarko and Schotter, in press).

remaining fraction α' of players are sophisticated like themselves.[45] Sophisticated players use the fEWA model to forecast what adaptive players will do, and choose strategies with high expected payoffs given their forecast and their guess about what sophisticated players will do. Denoting choice probabilities by adaptive and sophisticated players by $P_i^j(a,t)$ and $P_i^j(s,t)$, attractions for sophisticates are

$$A_i^j(s,t) \;=\; \sum_{k=1}^{m_{-i}} [\alpha' P_{-i}^k(s,t+1) + (1-\alpha') \cdot P_{-i}^k(a,t+1)] \cdot \pi_i(s_i^j, s_{-i}^k) \qquad (4.1)$$

Note that since the probability $P_{-i}^k(s,t+1)$ is derived from an analogous condition for $A_i^j(s,t)$, the system of equations is recursive. Self-awareness creates a whirlpool of recursive thinking which means QRE (and Nash equilibrium) are special cases in which all players are sophisticated and believe others are too ($\alpha = \alpha' = 1$).

An alternative structure we are currently studying links steps of sophistication to the steps of thinking used in the first period. For example, define zero learning steps as using fEWA; one step is best-responding to zero-step learners; two steps is best-responding to choices of one-step sophisticates, and so forth. We think this model can produce results similar to the recursive one we report below, and it replaces α and α' with τ from the theory of initial conditions so it reduces the entire thinking-learning-teaching model to only two parameters.

We estimate the sophisticated EWA model using data from p-beauty contests introduced above. Table 6 reports results and estimates of important parameters (with bootstrapped standard errors). For inexperienced subjects, adaptive EWA generates Cournot-like estimates ($\hat{\phi} = 0$ and $\hat{\delta} = .90$). Adding sophistication increases $\hat{\phi}$ and improves LL substantially both in- and out-of-sample. The estimated fraction of sophisticated players is 24% and their estimated perception $\hat{\alpha}'$ is zero, showing overconfidence (as in the thinking-steps estimates from the last section).[46]

Experienced subjects are those who play a second 10-period game with a different p parameter (the multiple of the average which creates the target number). Among

[45] To truncate the belief hierarchy, the sophisticated players believe that the other sophisticated players, like themselves, believe there are α' sophisticates.

[46] The gap between apparent sophistication and perceived sophistication shows the empirical advantage of separating the two. Using likelihood ratio tests, we can clearly reject both the rational expectations restriction $\alpha = \alpha'$ and the pure overconfidence restriction $\alpha' = 0$ although the differences in log-likelihood are not large.

Table 6: Sophisticated and adaptive learning model estimates for the p-beauty contest game (Camerer, Ho, and Chong, in press)

	inexperienced subjects		experienced subjects	
	sophisticated EWA	adaptive EWA	sophisticated EWA	adaptive EWA
ϕ	**0.44**	0.00	**0.29**	0.22
	$(0.05)^2$	(0.00)	(0.03)	(.02)
δ	**0.78**	0.90	**0.67**	0.99
	(0.08)	(0.05)	(0.05)	(0.02)
α	**0.24**	0.00	**0.77**	0.00
	(0.04)	(0.00)	(0.02)	(0.00)
α'	**0.00**	0.00	**0.41**	0.00
	(0.00)	(0.00)	(0.03)	(0.00)
LL				
(in sample)	**-2095.32**	-2155.09	**-1908.48**	-2128.88
(out of sample)	**-968.24**	-992.47	**-710.28**	-925.09

[2]Standard errors in parentheses.

experienced subjects, the estimated proportion of sophisticates increases to $\hat{\alpha} = 77\%$. Their estimated perceptions increase too but are still overconfident ($\hat{\alpha'} = 41\%$). The estimates reflect "learning about learning": Subjects who played one 10-period game come to realize an adaptive process is occurring; and most of them anticipate that others are learning when they play again.

4.2 Strategic teaching

Sophisticated players matched with the same players repeatedly often have an incentive to "teach" adaptive players, by choosing strategies with poor short-run payoffs which will change what adaptive players do, in a way that benefits the sophisticated player in the long-run. Game theorists have showed that strategic teaching could select one of many repeated-game equilibria (teachers will teach the pattern that benefits them) and could give rise to reputation formation without the complicated apparatus of Bayesian updating of Harsanyi-style payoff types (see Fudenberg and Levine, 1989; Watson, 1993; Watson and Battigali, 1997). This section of the paper describes a parametric model which embodies these intuitions, and tests it with experimental data. The goal is to show how the kinds of learning models described in the previous section can be parsimoniously extended to explain behavior in more complex games which are, perhaps, of even greater economic interest than games with random matching.

Consider a finitely-repeated trust game. A borrower B wants to borrow money from each of a series of lenders denoted L_i $(i = 1, \ldots, N)$. In each period a lender makes a single lending decision (*Loan* or *No Loan*). If the lender makes a loan, the borrower either (*repays* or *defaults*). The next lender in the sequence, who observed all the previous history, then makes a lending decision. The payoffs used in experiments are shown in Table 7.

There are actually two types of borrowers. As in post-Harsanyi game theory with incomplete information, types are expressed as differences in borrower payoffs which the borrowers know but the lenders do not (though the probability that a given borrower is each type is commonly known). The honest (Y) types actually receive *more* money from repaying the loan, an experimenter's way of inducing preferences like those of a person who has a social utility for being trustworthy (see Camerer, 2002, chapter 3 and references therein). The normal (X) types, however, earn 150 from defaulting and only

Table 7: Payoffs in the borrower-lender trust game, Camerer & Weigelt (1988)

lender strategy	borrower strategy	payoffs to lender	payoffs to borrower	
			normal (X)	honest (Y)
loan	default	-100	150	0
	repay	40	60	60
no loan	(no choice)	10	10	10

60 from repaying. If they were playing just once and wanted to earn the most money, they would default.

In the standard game-theoretic account, paying back loans in finite games arises because there is a small percentage of honest types who *always* repay. This gives normal-type borrowers an incentive to repay until close to the end, when they begin to use mixed strategies and default with increasing probability.

Whether people actually play these sequential equilibria is important to investigate for two reasons. First, the equilibria impose consistency between optimal behavior by borrowers and lenders and Bayesian updating of types by lenders (based on their knowledge and anticipation of the borrowers' strategy mixtures); whether reasoning or learning can generate this consistency is an open behavioral question (cf. Selten, 1978). Second, the equilibria are very sensitive to the probability of honesty (if it is too low the reputational equilibria disappear and borrowers should always default), and also make counterintuitive comparative statics predictions which are not confirmed in experiments (e.g., Neral and Ochs, 1992; Jung, Kagel and Levin, 1994).

In the experiments subjects play many sequences of 8 periods. The eight-period game is repeated to see whether equilibration occurs across many sequences of the entire game.[47] Surprisingly, the earliest experiments showed that the pattern of lending, default, and reactions to default across experimental periods within a sequence is roughly in line with the equilibrium predictions. Typical patterns in the data are shown in Figures 9a-b. Sequences are combined into ten-sequence blocks (denoted "sequence") and average frequencies are reported from those blocks. Periods 1,...,8 denote periods in each

[47]Borrower subjects do not play consecutive sequences, which removes their incentive to repay in the eighth period of one sequence so they can get more loans in the first period of the next sequence.

sequence. The figures show relative frequencies of no loan and default (conditional on a loan). Figure 9a shows that in early sequences lenders start by making loans in early periods (i.e., there is a low frequency of no-loan), but they rarely lend in periods 7-8. In later sequences they have learned to always lend in early periods and rarely lend in later periods. Figure 9b shows that borrowers rarely default in early periods, but usually default (conditional on getting a loan) in periods 7-8. The within-sequence pattern becomes sharper in later sequences.

The general patterns predicted by equilibrium are therefore present in the data. But given how complex the equilibrium is, how do players approximate it? Camerer and Weigelt (1988) concluded their paper as follows:

> ...the long period of disequilibrium behavior early in these experiments raises the important question of how people learn to play complicated games. The data could be fit to statistical learning models, though new experiments or new models might be needed to explain learning adequately. (pp 27-28)

The teaching model is a "new model" of the sort Camerer and Weigelt had in mind. It is a boundedly rational model of reputation formation in which the lenders learn whether to lend or not. They do not update borrowers' types and do not anticipate borrowers' future behavior (as in equilibrium models); they just learn.

In the teaching model, some proportion of borrowers are sophisticated and teach; the rest are adaptive and learn from experience but have no strategic foresight. The teachers choose strategies which are expected (given their beliefs about how borrowers will react to their teaching) to give the highest long-run payoffs in the remaining periods.

A sophisticated teaching borrower's attractions for sequence k after period t are specified as follows ($j \in \{repay, default\}$ is the borrower's set of strategies):

$$A_B^j(s,k,t) = \sum_{j'=Loan}^{NoLoan} P_L^{j'}(a,k,t+1)\cdot \pi_B(j,j') +$$

$$\max_{J_{t+1}}\{\sum_{v=t+2}^{T}\sum_{j'=Loan}^{NoLoan} \hat{P}_L^{j'}(a,k,v|j_{v-1}\in J_{t+1})\cdot \pi_B(j_v \in J_{t+1}, j')\}$$

The set J_{t+1} specifies a possible path of future actions by the sophisticated borrower from round $t+1$ until end of the game sequence. That is $J_{t+1} = \{j_{t+1}, j_{t+2}, \ldots, j_{T-1}, j_T\}$

and $j_{t+1} = j$.[48] The expressions $\hat{P}_L^{j'}(a, k, v|j_{v-1})$ are the overall probabilities of either getting a loan or not in the future periods v, which depends on what happened in the past (which the teacher anticipates).[49] $P_B^j(s, k, t+1)$ is derived from $A_B^j(s, k, t)$ using a logit rule.

The updating equations for adaptive players are the same as those used in fEWA with two twists. First, since lenders who play in later periods know what has happened earlier in a sequence, we assume that they learned from the experience they saw as if it had happened to them.[50]. Second, a lender who is about to make a decision in period 5 of sequence 17, for example, has *two* relevant sources of experience to draw on– the behavior she saw in periods 1-4 in sequence 17, *and* the behavior she has seen in the period 5's of the previous sequences (1-16). Since *both* kinds of experience could influence her current decision, we include both using a two-step procedure. After period 4 of sequence 17, for example, attractions for lending and not lending are first updated based on the period 4 experience. Then attractions are partially updated (using a degree of updating parameter σ) based on the experience in period 5 of the previous sequences.[51] The parameter σ is a measure of the strength of "peripheral vision"– glancing back at the "future" period 5's from previous sequences to help guess what lies ahead.

Of course, it is well-known that repeated-game behavior can arise in finite-horizon games when there are a small number of "unusual" types (who act like the horizon is unlimited), which creates an incentive for rational players to behave as if the horizon is unlimited until near the end (e.g., Kreps and Wilson, 1982). But specifying why some types are irrational, and how many they are, makes this interpretation difficult to test. In the teaching approach, which "unusual" type the teacher pretends to be arises endogenously from the payoff structure: They are Stackelberg types, who play the strategy they would choose if they could commit to it. For example, in trust games, they would like to commit to repaying; in entry-deterrence, they would like to commit

[48] To economize in computing, we search only paths of future actions that always have default following repay because the reverse behavior (repay following default) generates a lower return.

[49] Formally, $\hat{P}_L^{j'}(a,k,v|j_{v-1}) = \hat{P}_L^{Loan}(a,k,v-1|j_{v-1}) \cdot P_L^{j'}(a,k,v|(Loan, j_{v-1})) + \hat{P}_L^{NoLoan}(a,k,v-1|j_{v-1}) \cdot P_L^{j'}(a,k,v|(NoLoan, j_{v-1}))$.

[50] This is called "observational learning"; see Duffy and Feltovich, 1999) Without this assumption the model learns far slower than the lenders do so it is clear that they are learning from observing others

[51] The idea is to create an "interim" attraction for round t, $B_L^j(a,k,t)$, based on the attraction $A_L^j(a,k,t-1)$ and payoff from the round t, then incorporate experience in round $t+1$ from previous sequences, transforming $B_L^j(a,k,t)$ into a final attraction $A_L^j(a,k,t)$. See Camerer, Ho, and Chong (in press) for details.

to fighting entry.

The model is estimated using repeated game trust data from Camerer and Weigelt (1988). In Camerer, Ho and Chong (in press), we used parametric EWA to model behavior in trust games. That model allows two different sets of EWA parameters for lenders and borrowers. In this paper we use fEWA to model lenders and adaptive borrowers so the model has fewer parameters.[52] Maximum likelihood estimation is used to estimate parameters on 70% of the sequences in each experimental session, then behavior in the holdout sample of 30% of the sequences is forecasted using the estimated parameters.

As a benchmark alternative to the teaching model, we estimated an agent-based version of QRE suitable for extensive-form games (see McKelvey and Palfrey, 1998). Agent-QRE is a good benchmark because it incorporates the key features of repeated-game equilibrium– strategic foresight, accurate expectations about actions of other players, and Bayesian updating– but assumes stochastic best-response. We use an agent-based form in which players choose a distribution of strategies at each node, rather than using a distribution over all history-dependent strategies. We implement agent QRE with four parameters– different λ's for lenders, honest borrowers, and normal borrowers, and a fraction θ, the percentage of players with normal-type payoffs who are thought to act as if they are honest (reflecting a "homemade prior" which can differ from the prior induced by the experimental design[53]). (Standard equilibrium concepts are a special case of this model when λ's are large and $\theta = 0$, and fit much worse than AQRE does).

The models are estimated separately on each of the eight sessions to gauge cross-session stability. Since pooling sessions yields similar fits and parameter values, we report only those pooled results in Table 8 (excluding the λ values). The interesting parameters for sophisticated borrowers are estimated to be $\hat{\alpha} = .89$ and $\hat{\sigma} = .93$, which means most subjects are classified as teachers and they put a lot of weight on previous sequences. The teaching model fits in-sample and predicts better out-of-sample than AQRE by a modest margin (and does better in six of eight individual experimental sessions), predicting about 75% of the choices correctly. The AQRE fits reasonably well too (72% correct) but the estimated "homemade prior" θ is .91, which is absurdly high. (Earlier studies estimated numbers around .1-.2.) The model basically fits best by assuming that *all* borrowers

[52] We use four separate λ's, for honest borrowers, lenders, normal adaptive borrowers, and teaching borrowers, an initial attraction for lending $A(0)$, and the spillover parameter σ and teaching proportion α.

[53] see Camerer and Weigelt (1988), Palfrey and Rosenthal (1988), and McKelvey and Palfrey (1992).

Table 8: Model parameters and fit in repeated trust games

		model	
	statistic	fEWA+ teaching	Agent QRE
in-sample	hit rate (%)	76.5%	73.9%
calibration (n=5757)	log-likelihood	-2975	-3131
out-of-sample	hit rate (%)	75.8%	72.3%
validation (n=2894)	log-likelihood	-1468	-1544
parameters		estimates	
cross-sequence learning	σ	0.93	-
% of teachers	α	0.89	-
homemade prior p(honest)	θ	-	0.91

simply prefer to repay loans. This assumption fits most of the data but it mistakes teaching for a pure repayment preference. As a result, it does not predict the sharp upturn in defaults in periods 7-8, which the teaching model does.

Figures 5c-d show average predicted probabilities from the teaching model for the no-loan and conditional default rates. No-loan frequencies are predicted to start low and rise across periods, as they actually do, though the model underpredicts the no-loan rate in general. The model predicts the increase in default rate across periods reasonably well, except for underpredicting default in the last period.

The teaching approach as a boundedly-rational alternative to type-based equilibrium models of reputation-formation.[54] It has always seemed improbable that players are capable of the delicate balance of reasoning required to implement the type-based models, unless they learn the equilibrium through some adaptive process. The teaching model is one parametric model of that adaptive process. It retains the core idea in the theory of repeated games– namely, strategic foresight– and consequently, respects the fact that

[54]One direction we are pursuing is to find designs or tests which distinguish the teaching and equilibrium updating approaches. The sharpest test is to compare behavior in games with types that are fixed across sequences with types that are independently "refreshed" in each period within a sequence. The teaching approach predicts similar behavior in these two designs but type-updating approaches predict that reputation formation dissolves when types are refreshed.

matching protocols matter. And since the key behavioral parameters (α and σ) appear to be near one, restricting attention to these values should make the model workable for doing theory.

4.3 Summary

In this section we introduced the possibility that players can be sophisticated– i.e., they believe others are learning. (In future work, it would be interesting to link steps of iterated thinking, as in the first section, to steps of sophisticated thinking.) Sophistication links learning theories to equilibrium ones if sophisticated players are self-aware. Adding sophistication also improves the fit of data from repeated beauty-contest games. Interestingly, the proportion of estimated sophisticates is around a quarter when subjects are inexperienced, but rises to around three-quarters when the play an entire 10-period game a second time, as if subjects learn about learning. Sophisticated players who know they will be rematched repeatedly may have an incentive to "teach", which provides a boundedly rational theory of reputation formation. We apply this model to data on repeated trust games. The model adds only two behavioral parameters, representing the fraction of teachers and how much "peripheral vision" learners have (and some nuisance λ parameters), and predicts substantially better than a quantal response version of equilibrium.

5 Conclusion

In the introduction we stated that the research program in behavioral game theory has three goals: (1) To create a theory of one-shot or first-period play using an index of bounded rationality measuring steps of thinking; (2) to predict features of equilibration paths when games are repeated; and (3) to explain why players behave differently when matched together repeatedly.[55]

The models described in this paper illustrate ways to understand these three phenomena. There are lots of alternative models (especially of learning). The models described

[55] A fourth enterprise fits utility functions which reflect social preferences for fairness or equality. This important area is left aside in this paper.

here are just some examples of the style in which ideas can be expressed and how data are used to test and modify them.

Keep in mind that the goal is *not* to list deviations from Nash equilibrium and stop. Deviations are just hints. The goal is to develop alternative models which are precise, general, and disciplined by data. The thinking-steps model posits a Poisson distribution (with mean τ) of number of thinking steps, along with decision rules for what players using each number of steps will do. Studies with simple matrix games, beauty contests (unreported), mixed games, and entry games all show that values of τ around 1.5 can fit data reasonably well (and never worse than Nash equilibrium). The model is easy to use because players can be assumed to best-respond and the model usually makes realistic probabilistic predictions because the mixture of thinking steps types creates a population mixture of responses. The surprise is that the same model, which is tailor-made to produce spikes in dominance-solvable games, can also fit data from games with pure and mixed equilibria using roughly the same τ.

The second section compared several adaptive learning models. For explaining simple trends in equilibration, many of these models are close substitutes. However, it is useful to focus on where models fail if the goal is to improve them. The EWA hybrid was created to include the psychological intuitions behind both reinforcement learning (that received payoffs receive more weight than foregone payoffs) and belief learning (both types of payoffs receive equal weight). If both intuitions were compelling enough for people to want to compare them statistically, then a model which had both intuitions in it should be better still (and generally, it is). The functional fEWA uses one parameter (λ) and substitutes functions for parameters. The major surprise here is that functions like the change-detector $\phi_i(t)$ can reproduce differences across games in which parameter values fit best. This means the model can be applied to brand new games (when coupled with a thinking-steps theory of initial conditions) without having to make a prior judgment about what parameter values are reasonable, and without positing game-specific strategies. The interaction of learning and game structure creates reasonable parameter values automatically.

In the third section we extend the adaptive learning models to include sophisticated players who believe others are learning. Sophistication improves fit in the beauty contest game data (Experienced subjects seem to have "learned about learning" because the percentage of apparently sophisticated players is higher and convergence is faster.) Sophisticated players who realize they are matched with others repeatedly often have

an incentive to "teach" as in the theory of repeated games. Adding two parameters to adaptive learning was used to model learning and teaching in finitely-repeated trust games. While trustworthy behavior early in these games is known to be rationalizable by Bayesian-Nash models with "unusual" types, the teaching model create the unusual types from scratch. Teaching also fits and predicts better than more forgiving quantal response forms of the Bayesian-Nash type-based model. The surprise here is that the logic of mutual consistency and type updating is not needed to produce accurate predictions in finitely-repeated games with incomplete information.

5.1 Potential applications

A crucial question is whether behavioral game theory can help explain naturally-occurring phenomena. We conclude the paper with some speculations about the sorts of phenomena precise models of limited thinking, learning, and teaching could illuminate.

Bubbles: Limited iterated thinking is potentially important because, as Keynes and many others have pointed out before, it is not always optimal to behave rationally if you believe others are not. For example, prices of assets should equal their fundamental or intrinsic value if rationality is common knowledge (Tirole, 1985). But when the belief that others might be irrational arises, bubbles can too. Besides historical examples like Dutch tulip bulbs and the $5 trillion tech-stock bubble in the 1990s, experiments have shown such bubbles even in environments in which the asset's fundamental value is controlled and commonly-known.[56]

Speculation and competition neglect: The "Groucho Marx theorem" says that traders who are risk-averse should not speculate by trading with each other even if they have private information (since the only person who will trade with you may be better-informed). But this theorem rests on unrealistic assumptions of common knowledge of rationality and is violated constantly by massive speculative trading volume and other kinds of betting, as well as in experiments.[57]

Players who do limited iterated thinking, or believe others are not as smart as themselves, will neglect competition in business entry (see Camerer and Lovallo, 1999; Huber-

[56] See Smith, Suchanek and Williams, 1988; Camerer and Weigelt, 1993; and Lei, Noussair and Plott, 2001.

[57] See Sonsino, Erev and Gilat, 2000; Sovik, 2000.

man and Rubinstein, 2000). Competition neglect may partly explain why the failure rate of new businesses is so high. Managerial hubris, overconfidence, and self-serving biases which are correlated with costly delay and labor strikes in the lab (Babcock et al., 1995) and in the field (Babcock and Loewenstein, 1997) can also be interpreted as players not believing others always behave rationally.

Incentives: In a thorough review of empirical evidence on incentive contracts in organizations, Prendergast (1999) notes that workers typically react to simple incentives as standard models predict. However, firms usually do not implement complex contracts which *should* elicit higher effort and improve efficiency. Perhaps the firms' reluctance to bet on rational responses by workers is evidence of limited iterated thinking.

Macroeconomics: Woodford (2001) notes that in Phelps-Lucas "islands" models, nominal shocks can have real effects but their predicted persistence is too short compared to effects in data. He shows that imperfect information about *higher-order* nominal GDP estimates– beliefs about beliefs, and higher-order iterations– can cause longer persistence which matches the data. However, Svensson (2001) notes that iterated beliefs are probably constrained by computational capacity. If people have a projection bias, their beliefs about what others believe will be too much like their own, which undermines Woodford's case. On the other hand, in the thinking steps model players' beliefs are not mutually consistent so there is higher-order belief inconsistency which can explain longer persistence. In either case, knowing precisely how iterated beliefs work could help inform a central issue in macroeconomics– persistence of real effects of nominal shocks.

Learning: Other phenomena are evidence of a process of equilibration or learning. For example, institutions for matching medical residents and medical schools, and analogous matching in college sororities and college bowl games, developed over decades and often "unravel" so that high-quality matches occur before some agreed-upon date (Roth and Xing, 1994). Bidders in eBay auctions learn to bid late to hide their information about an object's common value (Bajari and Hortacsu, 2000). Consumers learn over time what products they like (Ho and Chong, 2000). Learning in financial markets can generate excess volatility and returns predictability, which are otherwise anomalous in rational expectations models (Timmerman, 1993). We are currently studying evolution of products in a high-uncertainty environment (electronics equipment) for which thinking-steps and learning models are proving useful.

Teaching: Teaching in repeated games may prove to be the most potentially useful tool

for economics, because it is essentially an account of how bounded rationality can give rise to some features of repeated-game behavior, where standard theory has been widely applied. The teaching model could be applied to repeated contracting, employment relationships, alliances among firms, industrial organization problems (such as pricing games among perennial rivals, and entry deterrence) and macroeconomic models of policymaker inflation-setting.[58]

References

[1] C. Anderson and C. F. Camerer, "Experience-weighted Attraction Learning in Sender-receiver Signaling Games," *Economic Theory*, 16(3), (2001), 689-718.

[2] J. Andreoni and J. Miller, "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence," *Economic Journal*, 103, (1993), 570-585.

[3] B. Arthur, "Designing Economic Agents That Act Like Human Agents: A Behavioral Approach to Bounded Rationality," *American Economic Review*, 81(2), (1991), 353-359.

[4] L. Babcock, G. Loewenstein, S. Issacharoff and C. F. Camerer, "Biased Judgments of Fairness in Bargaining," *American Economic Review*, 85, (1995), 1337-1343.

[5] Babcock, Linda and George Loewenstein. Explaining bargaining impasses: The role of self-serving biases. *Journal of Economic Perspectives*, 11 (1997), 109-126.

[6] P. Bajari and A. Hortacsu, "Winner's Curse, Reserve Prices, and Endogeneous Entry: Empirical Insights from e-Bay," Stanford University working paper, 2000.

[7] K. Binmore, J. Swierzbinski and C. Proulx, "Does Maximin Work? An Experimental Study," *Economic Journal*, 111, (2001), 445-464.

[8] R. Bloomfield, "Learning a Mixed Strategy Equilibrium in the Laboratory," *Journal of Economic Behavior and Organization*, 25, (1994), 411-436.

[58] We are currently applying the teaching model to the Kydland- Prescott model of commitment in which the public learns about inflation from past history (using the fEWA rule described below) and unemployment is determined by an expectational Phillips curve. Since policymakers face a temptation to choose surprisingly high inflation to lower unemployment, they can either act myopically or "teach" the public to expect low inflation which is Pareto-optimal in the long-run (cf. Sargent, 2000).

[9] A. Blume, D. DeJong, G. Neumann and N. Savin, "Learning in Sender-receiver Games," University of Iowa working paper, 1999.

[10] G. Brown, "Iterative Solution of Games by Fictitious Play," in *Activity Analysis of Production and Allocation*, John Wiley & Sons, New York, 1951.

[11] R. Bush and F. Mosteller, *Stochastic Models for Learning*, John Wiley & Sons, New York, 1955.

[12] A. Cabrales and W. Garcia-Fontes, "Estimating Learning Models with Experimental Data," University of Pompeu Febra working paper 501, 2000.

[13] G. P. Cachon and C. F. Camerer, "Loss-avoidance and Forward Induction in Experimental Coordination Games," *The Quarterly Journal of Economics*, 111, (1996), 165-194.

[14] C. F. Camerer, "Do Biases in Probability Judgment Matter in Markets? Experimental Evidence," *American Economic Review*, 77, (1987), 981-997.

[15] C. F. Camerer, "Progress in Behavioral Game Theory," *Journal of Economic Perspectives*, 11 (1997), 167-188.

[16] C. F. Camerer, *Behavioral Game Theory: Experiments on Strategic Interaction*, Princeton: Princeton University Press, 2002. http://www.hss.caltech.edu/CourseSites/Psy101/psy101.html

[17] C.F. Camerer and T-H Ho, "EWA Learning in Normal-form Games: Probability Rules, Heterogeneity and Time Variation," *Journal of Mathematical Psychology*, 42 (1998), 305-326.

[18] C. F. Camerer and T-H Ho, "Experience-weighted Attraction Learning in Normal-form Games," *Econometrica*, 67 (1999), 827-874.

[19] C. F. Camerer and T-H Ho, "Strategic Learning and Teaching in Games," in S. Hoch and H. Kunreuther, eds., *Wharton on Decision Making*, New York: Wiley, 2001.

[20] C. F. Camerer, T-H Ho and J. K. Chong, "Sophisticated EWA Learning and Strategic Teaching in Repeated Games," *Journal of Economic Theory*, in press. http://www.hss.caltech.edu/c̃amerer/camerer.html

[21] C. F. Camerer, T-H Ho, and X. Wang, "Individual Differences and Payoff Learning in Games," University of Pennsylvania working paper, 1999.

[22] C. F. Camerer, D. Hsia, and T-H Ho, "EWA Learning in Bilateral Call Markets," in *Experimental Business Research*, ed. by A. Rapoport and R. Zwick, in press.

[23] C. F. Camerer, E. Johnson, S. Sen and T. Rymon, "Cognition and Framing in Sequential Bargaining for Gains and Losses," in *Frontiers of Game Theory*, ed. by K Binmore, A. Kirman, and P. Tani., MIT Press, Cambridge, (1993), 27-48.

[24] C. F. Camerer and D. Lovallo, "Overconfidence and Excess Entry: An Experimental Approach," *American Economic Review*, 89, (1999), 306-318.

[25] C. F. Camerer and K. Weigelt, "Experimental Tests of A Sequential Equilibrium Reputation Model," *Econometrica*, 56, (1988), 1-36.

[26] C. F. Camerer and K. Weigelt, "Convergence in Experimental Double Auctions for Stochastically Lived Assets," in D. Friedman and J. Rust (Eds.), *The Double Auction Market: Theories, Institutions and Experimental Evaluations*, Redwood City, CA: Addison-Wesley, (1993), 355-396.

[27] M. Capra, "Noisy Expectation Formation in One-shot Games," Unpublished dissertation, University of Virginia, 1999.

[28] M. Capra, J. Goeree, R. Gomez and C. Holt, "Anomalous Behavior in a Traveler's Dilemma," *American Economic Review*, 89, (1999), 678-690.

[29] Y-W Cheung and D. Friedman, "Individual Learning in Normal Form Games: Some Laboratory Results," *Games and Economic Behavior*, 19, (1997), 46-76.

[30] M. H. Christiansen and N. Chater, "Toward a Connectionist Model of Recursion in Human Linguistic Performance," *Cognitive Science*, 23, (1991), 157-205.

[31] K. Clark and M. Sefton, "Matching Protocols in Experimental Games," University of Manchester working paper, 1999.

[32] D. Cooper, and J. Van Huyck, "Evidence on the Equivalence of the Strategic and Extensive Form Representation of Games," Texas A&M University Department of Economics, 2001. http://econlab10.tamu.edu/JVH_gtee/

[33] M. Costa-Gomes, V. Crawford and B. Broseta, "Cognition and Behavior in Normal-form Games: An Experimental Study," *Econometrica*, 69, (2001), 1193-1235.

[34] V. Crawford, "Theory and Experiment in the Analysis of Strategic Interactions," in D. Kreps and K. Wallis (Eds.), *Advances in Economics and Econometrics: Theory and Applications*, Seventh World Congress, Volume I. Cambridge: Cambridge University Press, 1997.

[35] F. Crick, *What Mad Pursuit?*, 1988.

[36] R. T. A. Croson, "Thinking Like A Game Theorist: Factors Affecting the Frequency of Equilibrium Play," *Journal Of Economic Behavior And Organization*, 41(3), (2000), 299-314.

[37] J. Cross, *A Theory of Adaptive Learning Economic Behavior*, New York: Cambridge University Press, 1983.

[38] J. Duffy and N. Feltovich, "Does Observation of Others Affect Learning in Strategic Environments? An Experimental Study," *International Journal of Game Theory*, 28, (1999), 131-152.

[39] I. Erev, Y. Bereby-Meyer and A. Roth, "The Effect of Adding a Constant to All Payoffs: Experimental Investigation, and a Reinforcement Learning Model with Self-Adjusting Speed of Learning," *Journal of Economic Behavior and Organization*, 39, (1999), 111-128.

[40] I. Erev and A. Roth, "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed-strategy Equilibria," *The American Economic Review*, 88 (1998), 848-881.

[41] D. Fudenberg and D. Levine, "Reputation and Equilibrium Selection in Games with A Patient Player," *Econometrica*, 57 (1989), 759-778.

[42] D. Fudenberg and D. Levine, *The Theory of Learning in Games*, Boston: MIT Press, 1998.

[43] J. K. Goeree and C. A. Holt, "Stochastic Game Theory: For Playing Games, Not Just for Doing Theory," *Proceedings of the National Academy of Sciences*, 96, (1999a), 10564-10567.

[44] J. K. Goeree and C. A. Holt, " A Theory of Noisy Introspection," University of Virginia Department of Economics, 1999b.

[45] J. K. Goeree and C. A. Holt, "Ten little treasures of game theory, and ten contradictions," *American Economic Review*, in press.

[46] C. Harley, "Learning the Evolutionary Stable Strategies," *Journal of Theoretical Biology*, 89, (1981), 611-633.

[47] D. Heller and R. Sarin, "Parametric Adaptive Learning," University of Chicago Working Paper, 2000.

[48] T-H Ho, C. F. Camerer, and J. K. Chong, "Economic Value of EWA Lite: A Functional Theory of Learning in Games," University of Pennsylvania working paper, 2001. http://www.hss.caltech.edu/ camerer/camerer.html

[49] T-H Ho, C. F. Camerer and K. Weigelt, "Iterated Dominance and Iterated Best Response in Experimental "p-Beauty Contests," *American Economic Review*, 88, (1998), 947-969.

[50] T-H Ho and J. K. Chong, "A Parsimonious Model of SKU Choice," University of Pennsylvania working paper, 1999.

[51] T-H Ho and K. Weigelt, "Task Complexity, Equilibrium Selection, and Learning: An Experimental Study," *Management Science*, 42, (1996), 659-679.

[52] E. Hopkins, "Two Competing Models of how People Learn in Games," *Econometrica*, in press.

[53] G. Huberman and A. Rubinstein, "Correct Belief, Wrong Action and a Puzzling Gender Difference," working paper 2000.

[54] E. J. Johnson, C. F. Camerer, S. Sen and T. Rymon, "Detecting Backward Induction in Sequential Bargaining Experiments," *Journal of Economic Theory*, in press. http://www.hss.caltech.edu/ camerer/camerer.html

[55] Y. J. Jung, J. H. Kagel, and D. Levin, "On the Existence of Predatory Pricing: An Experimental Study of Reputation and Entry Deterrence in the Chain-store Game," *RAND Journal of Economics*, 25, (1994), 72-93.

[56] D. Kahneman, "Experimental Economics: A Psychological Perspective," in R. Tietz, W. Albers, and R. Selten (Eds.) *Bounded Rational Behavior in Experimental Games and Markets*, New York: Springer-Verlag, 1988, 11-18.

[57] H. Kaufman and G. M. Becker, "The Empirical Determination of Game- theoretical Strategies," *Journal of Experimental Psychology*, 61, (1961), 462-468.

[58] M. G. Kocher and M. Sutter, "When the 'Decision Maker Matters: Individual versus Team Behavior in Experimental 'Beauty-contest Games," University of Innsbruck Institute of Public Economics discussion paper 2000/4, 2000.

[59] D. Kreps and R. Wilson, "Reputation and Imperfect Information," *Journal of Economic Theory*, 27 (1982), 253-279.

[60] V. Lei, C. Noussair and C. Plott, "Non-speculative Bubbles in Experimental Asset Markets: Lack of Common Knowledge of Rationality," *Econometrica*, 69 (2001), 813-59.

[61] R. G. Lucas, "Adaptive Behavior and Economic Theory," *Journal of Business*, 59 (October 1986), S401-S426.

[62] G. Mailath, "Do People Play Nash Equilibrium? Lessons from Evolutionary Game Theory," *Journal of Economic Literature*, 36, (1998), 1347-1374.

[63] D. Malcolm and B. Lieberman, "The Behavior of Responsive Individuals Playing a Two-person Zero-sum Game Requiring the Use of Mixed Strategies," *Psychonomic Science*, 2, (1965), 373-374.

[64] P. H. McAllister, "Adaptive Approaches to Stochastic Programming," *Annals of Operations Research*, 30, (1991), 45-62.

[65] R. D. McKelvey and T. R. Palfrey, "An Experimental Study of the Centipede Game," *Econometrica*, 60, (1992), 803-836.

[66] R. D. McKelvey and T. R. Palfrey, "Quantal Response Equilibria for Normal-form Games," *Games and Economic Behavior*, 7, (1995), 6-38.

[67] R. D. McKelvey and T. R. Palfrey, "Quantal Response Equilibria for Extensive-form Games," *Experimental Economics*, 1, (1998), 9-41.

[68] D. Mookerjee and B. Sopher, "Learning Behavior in an Experimental Matching Pennies Game," *Games and Economic Behavior*, 7, (1994), 62-91.

[69] D. Mookerjee and B. Sopher, "Learning and Decision Costs in Experimental Constant-sum Games," *Games and Economic Behavior*, 19, (1997), 97-132.

[70] R. Nagel, "Experimental Results on Interactive Competitive Guessing," *The American Economic Review*, 85, (1995), 1313-1326.

[71] R. Nagel, "A Review of Beauty Contest Games," in *Games and Human Behavior: Essays in honor of Amnon Rapoport*, ed. by D. Budescu, I. Erev, and R. Zwick., Lawrence Erlbaum Assoc. Inc., New Jersey, 1999, 105-142.

[72] R. Nagel, A. Bosch-Domenech, A. Satorra and J. Garcia-Montalvo, "One, Two, (Three), Infinity: Newspaper and Lab Beauty-contest Experiments," Universitat Pompeu Fabra, 1999.

[73] J. Neral and J. Ochs, "The Sequential Equilibrium Theory of Reputation Building: A Further Test," *Econometrica*, 60, (1992), 1151-1169.

[74] Y. Nyarko and A. Schotter, "An Experimental Study of Belief Learning Using Elicited Beliefs," *Econometrica*, in press.

[75] J. Ochs, "Games with Unique, Mixed Strategy Equilibria: An Experimental Study," *Games and Economic Behavior*, 10, (1995), 202-217.

[76] J. Ochs, "Entry in Experimental Market Games," in *Games and Human Behavior: Essays in honor of Amnon Rapoport*, ed. by D. Budescu, I. Erev, and R. Zwick., Lawrence Erlbaum Assoc. Inc., New Jersey, 1999.

[77] B. O'Neill, "Nonmetric Test of the Minimax Theory of Two-person Zero-sum Games," *Proceedings of the National Academy of Sciences*, 84, (1987), 2106-2109.

[78] T. R. Palfrey and H. Rosenthal, "Private Incentives in Social Dilemmas: The Effects of Incomplete Information and Altruism," *Journal of Public Economics*, 35 (1988), 309-332.

[79] J. Partow and A. Schotter, "Does Game Theory Predict Well for the Wrong Reasons? An Experimental Investigation," C.V. Starr Center for Applied Economics working paper 93-46, New York University, 1993.

[80] C. R. Plott, "Dimensions of Parallelism: Some Policy Applications of Experimental Methods," in A. E. Roth (ed.), *Laboratory experimentation in economics: Six points of view*, Cambridge: Cambridge University Press, 1986.

[81] C. R. Plott, "Laboratory Experimental Testbeds: Applications to the PCS Auctions," *Journal of Economics and Management Strategy*, 6, (1997), 605-638.

[82] C. Prendergast, "The Provision of Incentives in Firms," *Journal of Economic Literature*, 37, (March 1999), 7-63.

[83] A. Rapoport and W. Amaldoss, "Mixed Strategies and Iterative Elimination of Strongly Dominated Strategies: An Experimental Investigation of States of Knowledge," *Journal of Economic Behavior and Organization*, 42, (2000), 483-521.

[84] A. Rapoport and R. B. Boebel, Richard B, "Mixed Strategies in Strictly Competitive Games: A Further Test of the Minimax Hypothesis," *Games and Economic Behavior*, 4, (1992), 261-283.

[85] A. Rapoport and I. Erev, "Coordination, "Magic", and Reinforcement Learning in a Market Entry Game," *Games and Economic Behavior*, 23, (1998), 146-175.

[86] A. Rapoport, A. K-C Lo, and R. Zwick, "Choice of Prizes Allocated by Multiple Lotteries with Endogenously Determined Probabilities," University of Arizona, Department of Management and Policy working paper, 1999.

[87] S. J. Rassenti, V. L. Smith and B. J. Wilson, "Turning off the Lights," *Regulation*, Fall 2001, 70-76.

[88] A. Roth, "The Economist as Engineer," *Econometrica*, in press.

[89] A. Roth, G. Barron, I. Erev and R. Slonim, "Equilibrium and Learning in Economic Environments: the Predictive Value of Approximations," Harvard University Working Paper, 2000.

[90] A. Roth and I. Erev, "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term," *Games and Economic Behavior*, 8, (1995), 164-212.

[91] A. Roth and X. Xing, "Jumping the Gun: Imperfections and Institutions Related to the Timing of Market Transactions," *American Economic Review*, 84, (1994), 992-1044.

[92] T. Salmon, "Evidence for 'Learning to Learn' Behavior in Normal-form Games," Caltech working paper, 1999.

[93] T. Salmon, "An Evaluation of Econometric Models of Adaptive Learning," *Econometrica*, 69, (2001), 1597-1628.

[94] T. Sargent, *The Conquest of American Inflation*, Princeton: Princeton University Press, 1999.

[95] T. Schelling, *The Strategy of Conflict*, Harvard University Press, 1960.

[96] A. Schotter, K. Weigelt and C. Wilson, "A laboratory investigation of multiperson rationality and presentation effects," *Games and Economic Behavior*, 6, (1994), 445-468.

[97] D. A. Seale and A. Rapoport, "Elicitation of Strategy Profiles in Large Group Coordination Games," *Experimental Economics*, 3, (2000), 153-179.

[98] R. Selten, "The Chain Store Paradox," *Theory and Decision*, 9, (1978), 127-159.

[99] R. Selten and R. Stoecker, "End Behavior in Sequences of Finite Prisoner's Dilemma Supergames: A Learning Theory Approach," *Journal of Economic Behavior and Organization*, 7 (1986), 47-70.

[100] V. L. Smith, G. Suchanek and A. Williams, "Bubbles, Crashes and Endogeneous Expectations in Experimental Spot Asset Markets," *Econometrica*, 56, (1988), 1119-1151.

[101] D. Sonsino, I. Erev and S. Gilat, "On the Likelihood of Repeated Zero-sum Betting by Adaptive (Human) Agents," Technion, Israel Institute of Technology, 2000.

[102] Y. Sovik, "Impossible Bets: An Experimental Study," University of Oslo Department of Economics, 1999.

[103] D. O. Stahl, "Boundedly Rational Rule Learning in a Guessing Game," *Games and Economic Behavior*, 16 (1996), 303-330.

[104] Stahl, Dale O., "Sophisticated Learning and Learning Sophistication," University of Texas at Austin Working Paper, 1999.

[105] D. O. Stahl, "Local Rule Learning in Symmetric Normal-form Games: Theory and Evidence," *Games and Economic Behavior*, 32 (2000), 105-138.

[106] D. O. Stahl, and P. Wilson, "On Players Models of Other Players: Theory and Experimental Evidence," *Games and Economic Behavior*, 10 (1995), 213-54.

[107] J. A. Sundali, A. Rapoport and D. A. Seale, "Coordination in Market Entry Games with Symmetric Players," *Organizational Behavior and Human Decision Processes*, 64, (1995), 203-218.

[108] R. Sutton and A. Barto, *Reinforcement Learning: An Introduction*, Boston: MIT Press, 1998.

[109] L. E. O. Svensson, "Comments on Michael Woodford paper," presented at Knowledge, Information and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps. Columbia University, October 5-6, 2001.

[110] F.-F. Tang, "Anticipatory Learning in Two-person Games: Some Experimental Results," *Journal of Economic Behavior and Organization*, 44 (2001), 221-232.

[111] A. G. Timmerman, "How Learning in Financial Markets Generates Excess Volatility and Predictability in Stock Prices," *Quarterly Journal of Economics*, 108 (November 1993) 1135-1145.

[112] J. Tirole, "Asset Bubbles and Overlapping Generations," *Econometrica*, 53, (1985), 1071-1100 (reprinted 1499-1528).

[113] E. Van Damme, "Game Theory: The Next Stage," in L. A. Gerard-Varet, Alan P. Kirman, and M. Ruggiero (Eds.), *Economics beyond the Millennium*, Oxford University Press, 1999, 184-214.

[114] J. Van Huyck, R. Battalio, and R. Beil, "Tacit Cooperation Games, Strategic Uncertainty, and Coordination Failure," *The American Economic Review*, 80, (1990), 234-248.

[115] J. Van Huyck, R. Battalio, and F. W. Rankin. "Selection Dynamics and Adaptive Behavior Without Much Information," Texas A & M Department of Economics working paper, 2001.

[116] J. Van Huyck, J. Cook and R. Battalio, "Adaptive Behavior and Coordination Failure," *Journal of Economic Behavior and Organization*, 32, (1997), 483-503.

[117] Von Neumann and Morgenstern, *The Theory of Games and Economic Behavior*, Princeton: Princeton University Press, 1944.

[118] M. Warglien, M. G. Devetag and P. Legrenzi, "Mental Models and Nave Play in Normal Form Games," Universita Ca' Foscari di Venezia working paper, 1998.

[119] J. Watson, "A 'Reputation' Refinement without Equilibrium," *Econometrica*, 61, (1993), 199-205.

[120] J. Watson and P. Battigali, "On 'Reputation' Refinements with Heterogeneous Beliefs," *Econometrica*, 65, (1997), 363-374.

[121] J. Weibull, *Evolutionary Game Theory*, Cambridge: MIT Press, 1995.

[122] G. Weizsacker, "Ignoring the Rationality of Others: Evidence from Experimental Normal-form Games," Harvard Business School working paper, 2000.

[123] M. Woodford, "Imperfect common knowledge and the effects of monetary policy," presented at Knowledge, Information and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps. Columbia University, October 5-6, 2001.

6 Appendix: Thinking models applied to mixed games and entry games

6.1 Games with mixed equilibria

A good model of thinking steps should be able to *both* account for deviations from Nash equilibrium (as in the games above), *and* reproduce the successes of Nash equilibrium. A domain in which Nash equilibrium does a surprisingly good job is in games with unique mixed equilibria. It is hard to beat Nash equilibrium in these games because (as we shall see) the correlation with data is actually very good (around .9) so there is little room for improvement. Instead, the challenge is to see how well a thinking-steps model which bears little resemblance to the algebraic logic of equilibrium mixing can approximate behavior in these games.

Early tests in the 1960's and 1970's (mostly by psychologists) appeared to reject Nash equilibrium as a description of play in mixed games. As others have noted (e.g., Binmore

et al., 2001), these experiments were incomplete in important dimensions and hence inconclusive. Financial incentives were very low or absent; subjects typically did not play other human subjects (and often were deceived about playing other people, or were only vaguely instructed about how their computer opponents played); and pairs were often matched repeatedly so that (perceived) detection of temporal patterns permitted subjects to choose nonequilibrium strategies. Under conditions ideal for equilibration, however, convergence was rapid and sharp. Kaufman and Becker (1961), for example, had subjects specify mixtures and told them that a computer program would then choose a mixture to minimize the subjects' earnings. Subjects could maximize their possible gains by choosing the Nash mixture. After playing five games, more than half learned to do so. More recent experiments are also surprisingly supportive of Nash equilibrium (see Binmore et al., 2001; and Camerer, 2002, chapter 2). The data are supportive in two senses- (1) equilibrium predictions and actual frequencies are closely correlated, when taken as a whole (e.g., strategies predicted to be more likely are almost always played more often); and (2) it is hard to imagine any parsimonious theory which can explain the modest deviations.

We applied a version of the thinking model in which $K-$step thinkers think all others are using $K-1$ steps along with best response to see whether it could produce predictions as accurate as Nash in games with mixed equilibria. This model is extremely easy to use (just start with step zero mixtures and compute best responses iteratively). Furthermore, it creates natural "purification": Players using different thinking steps usually choose pure strategies, but the Poisson distribution of steps generates a mixture of responses and hence, a probabilistic prediction.

Model predictions are compared with data from 15 games with unique mixed equilibria reported in Camerer (2002, chapter 2).[59] These games are not a random or exhaustive sample of recent research but there are enough observations that we are confident the basic conclusion will be overturned by adding more studies. Note that we use data from *all* the periods of these games rather than the first period only. (In most cases the first-period data are rarely reported, and there is usually little trend over time in the data.)

[59]The studies, in the order in which τ estimates are reported below, are Malcolm and Lieberman (1965), O'Neill (1987), Rapoport and Boebel (1992), Mookerjhee and Sopher (1997), Tang (2001, games 3 and 1), Ochs (1995, games with 9 and 4 payoffs), Bloomfield (1994), Rapoport and Amaldoss (2000, r=8, 20), Binmore, Swierzbinski, and Proulx (2001), games 1, 3, 4. Readers should let us know of published studies which we overlooked and we'll plan to include them in a later draft.

Each data point in Figure 11 represents a single strategy from a different game (pooling across all periods to reduce sampling error).[60] Figure 11 plots actual frequencies on the ordinate (y) axis against either mixed-strategy equilibrium predictions or thinking-steps predictions on the abscissa (x) axis.

In Figure 11, the value of τ is common across games (1.46) and minimizes mean squared deviations between predicted and actual frequencies. When values are estimated separately for each game to minimize mean squared deviations, the values across the games (in the order they are listed above) are .3, .3, .3, 2.2, 2.5, .1, 1.8, 2.3, 2.9, 2.7, .5, .8, 1.6, 1.5, 1.9. The lower values occur in games where the actual mixtures are close to equal across strategies, so that a distribution with $\tau = 0$ fits well. When there are dominated strategies, which are usually rarely played, much higher τ values are need since low τ generates a lot of random play and frequent dominance violation. The simple arithmetic average across the 15 games is 1.45 which is very close to the best-fitting common $\tau = 1.46$. The Figure shows two regularities: Both thinking-steps (circles in the plot) and equilibrium predictions (triangles) have very high correlations (above .9) with the data, though there is substantial scatter, especially at high probabilities. (Keep in mind that sampling error means there is an upper bound on how well any model could fit- even the true model which generated the data.) The square roots of the mean squared deviation is around .10 for both models.

While the thinking-steps model with common τ is a little less accurate than Nash equilibrium (the game-specific model is more accurate), the key point is that the same model which can explain Nash deviations in dominance-solvable games and matrix games fits about as well with a value of τ close to those estimated in other games.

Table 9 shows a concrete example of how the thinking model is able to approximate mixture probabilities. The game is Mookerjhee and Sopher's (1997) 4x4 game. Payoffs are shown as wins (+) or losses (-) (the (2/3)+ means a 2/3 chance of winning) in the upper left cells. The rightmost columns show the probabilities with which row players using different thinking steps choose each of the four row strategies. To narrate a bit, zero-step players randomize (each is played with probability .25); one-step players best-respond to a random column choice and choose row strategy 3 with probability 1, and so forth. First notice that the weakly dominated strategy (4) is only chosen by a quarter of 0-step

[60] In each game, data from those n-1 out of the n possible strategies with the most extreme predicted equilibrium probabilities are used to fit the models. Excluding the n-th strategy reduces the dependence among data points, since all n frequencies (and predictions) obviously add to one.

Table 9: How thinking steps fits mixed-game data (Mookerjhee and Sopher 4x4 game)

row		strategies 1-4			column thinking steps 0-5								thinking
strat's	1	2	3	4	step 0	1	2	3	4	5	MSE	data	model
1	+	-	-	+	0.25	0	0.5	1	0	0	0.375	.32	.37
2	-	-	+	+	0.25	0	0	0	1	0	0.25	.17	.14
3	-	+	(2/3)+	(2/3)+	0.25	1	0.5	0	0	1	0.375	.43	.46
4	-	-	(1/3)+	+	0.25	0	0	0	0	0	0	.08	.03
step 0	.25	.25	.25	.25									
1	0.5	0.5	0	0									
2	1	0	0	0									
3	0	0	1	0									
4	0	0.5	0.5	0									
5	0.5	0.5	0	0									
6	1	0	0	0									
7	0	0	1	0									
MSE	0.375	0.25	0.375	0									
data	0.38	0.31	0.27	0.04									
thinking	0.46	0.23	0.28	0.03									

players (since it is never a best response against players who randomize), which generates a small amount of choice that matches the data. Notice also that the best responses tend to lurch around as thinking steps change. When these are averaged over thinking steps frequencies a population mixture results. Furthermore, one of the quirkiest features of mixed equilibrium is that one player's mixture depends only on the other player's payoffs. This effect also occurs in the thinking steps models because a K-step row player's payoffs affect row's best responses, which then affect a $K+1$-step column player's best response. So one player's payoffs affect the other's strategies indirectly. The table also shows the MSE (mixed equilibrium) prediction, the data frequencies, and overall frequencies from the thinking-steps model when $\tau = 2.2$. The model fits the data closer than MSE for row players (it accounts for underplay of row strategy 2 and overplay of strategy 3) and is equally accurate as MSE for column players. As noted in the text, the point is not that the thinking steps model outpredicts MSE– it cannot, because MSE has such a high correlation– but simply that the same model which explains behavior in dominance-solvable, matrix, and entry games also generates mixtures of players playing near-pure strategies which are close to outcomes in mixed games.

6.2 Market entry games

Analysis of the simple entry game described in the text proceeds as follows. Step 0's randomize so $f(0)/2$ level 0's enter. Define the relative proportion of entry after accounting up through level k as $N(k)$. Define a Boolean function $B(X) = 1$ if X true, $B(X) = 0$ if X false. Level 1's enter iff $1/2 < c$. Therefore, total entry "after" level 1 types are accounted for is $N(1) = f(0)/2 + B[N(0)/(f(0)/2) > c]f(1)$ Total entry after level k type is therefore $N(k) = f(0)/2 + \sum_{n=1}^{k} f(n)B[N(n-1)/(\sum_{m=1}^{k} f(m)) > c]$. A given c and τ then generates a sequence of cumulated entry rates which asymptotes as k grows large. Define a function $N(all|\tau)(c)$ as the overall rate of entry, given τ, for various capacities c.

The data reported in the text Figure come from experiments by Sundali et al. (1995) and Seale and Rapoport (1999). Their game is not precisely the same as the one analyzed because in their game entrants earn $1 + 2(c - e)$ (where e is the number of entrants) and nonentrants earn 1. They used 20 subjects with odd values of c (1,3,...19). To compute entry rates reported in the Figure we averaged entry for adjacent c values (i.e., averaging 1 and 3 yields a figure for $c = 2$ to match c=.1, averaging 3 and 5 yields a figure for

$c = 4$ corresponding to c=.2, etc.) Obviously the analysis and data are not perfectly matched, but we conjecture that the thinking steps model can still match data closely and reproduce the three experimental regularities described in the text; whether this is true is the subject of ongoing research.

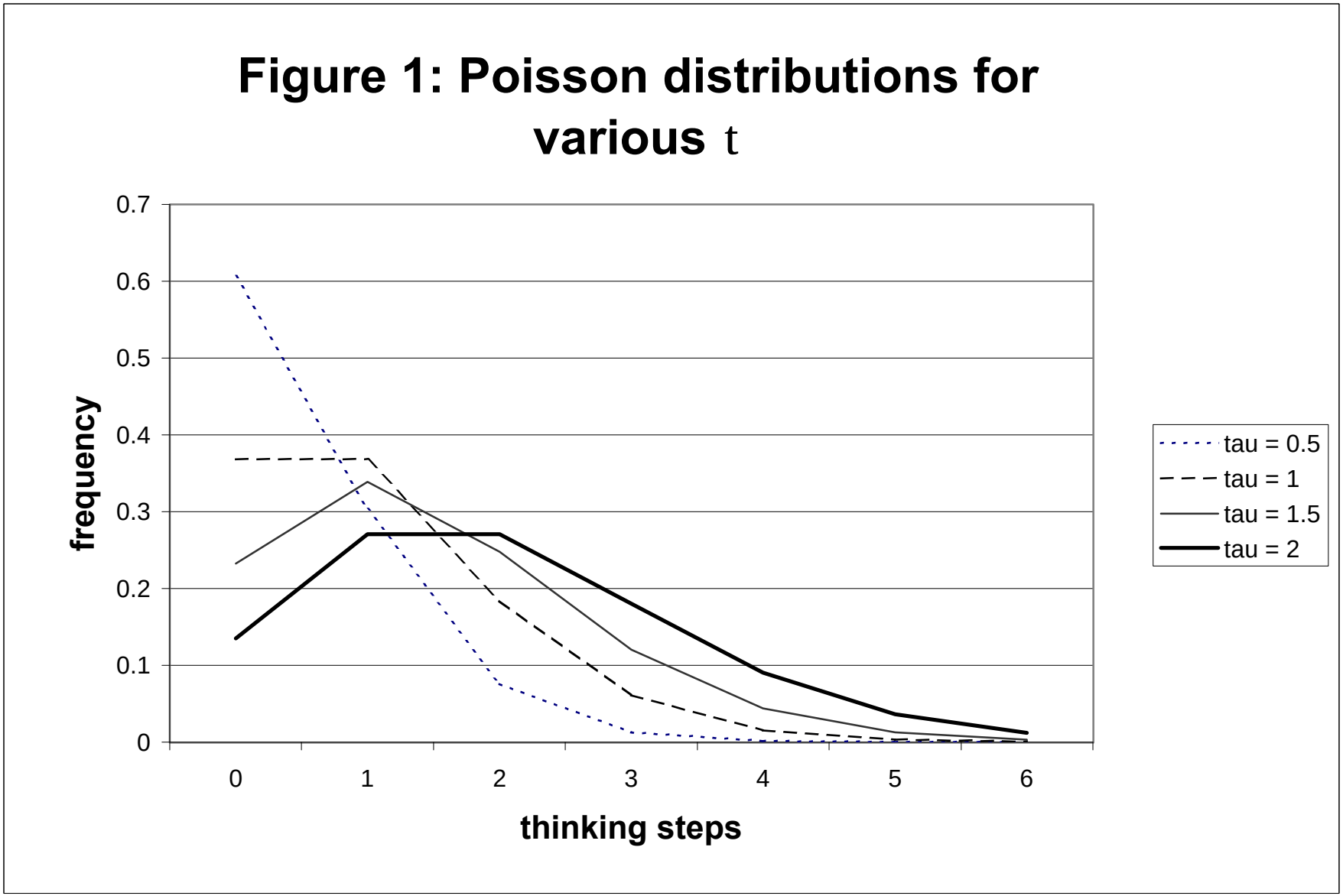


Figure 1: Poisson distributions for various t
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
frequency
0
1
2
3
4
5
6
thinking steps
tau = 0.5
tau = 1
tau = 1.5
tau = 2

Figure 2: Fit of thinking-steps model to three games (R^2=.84)

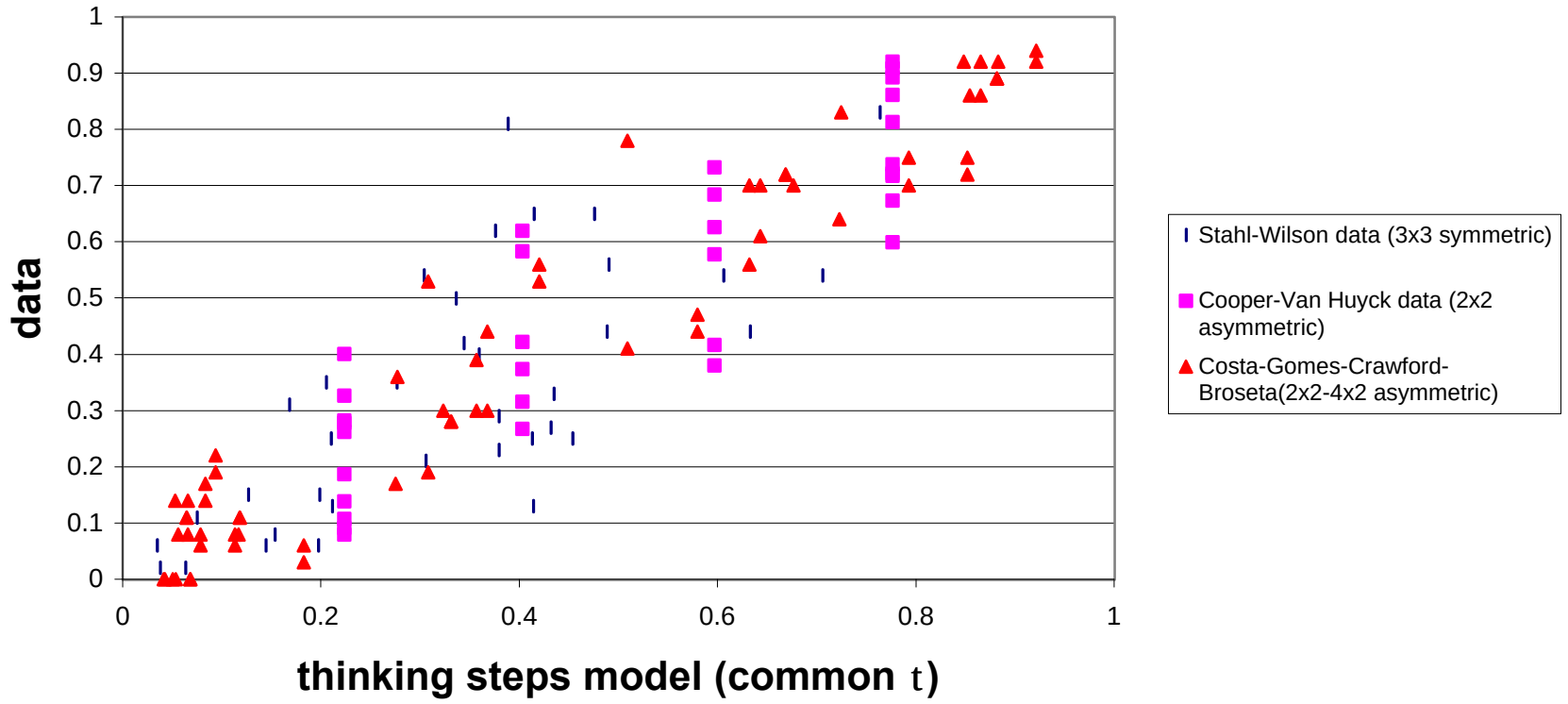


Figure 3: Nash equilibrium predictions vs data in three games

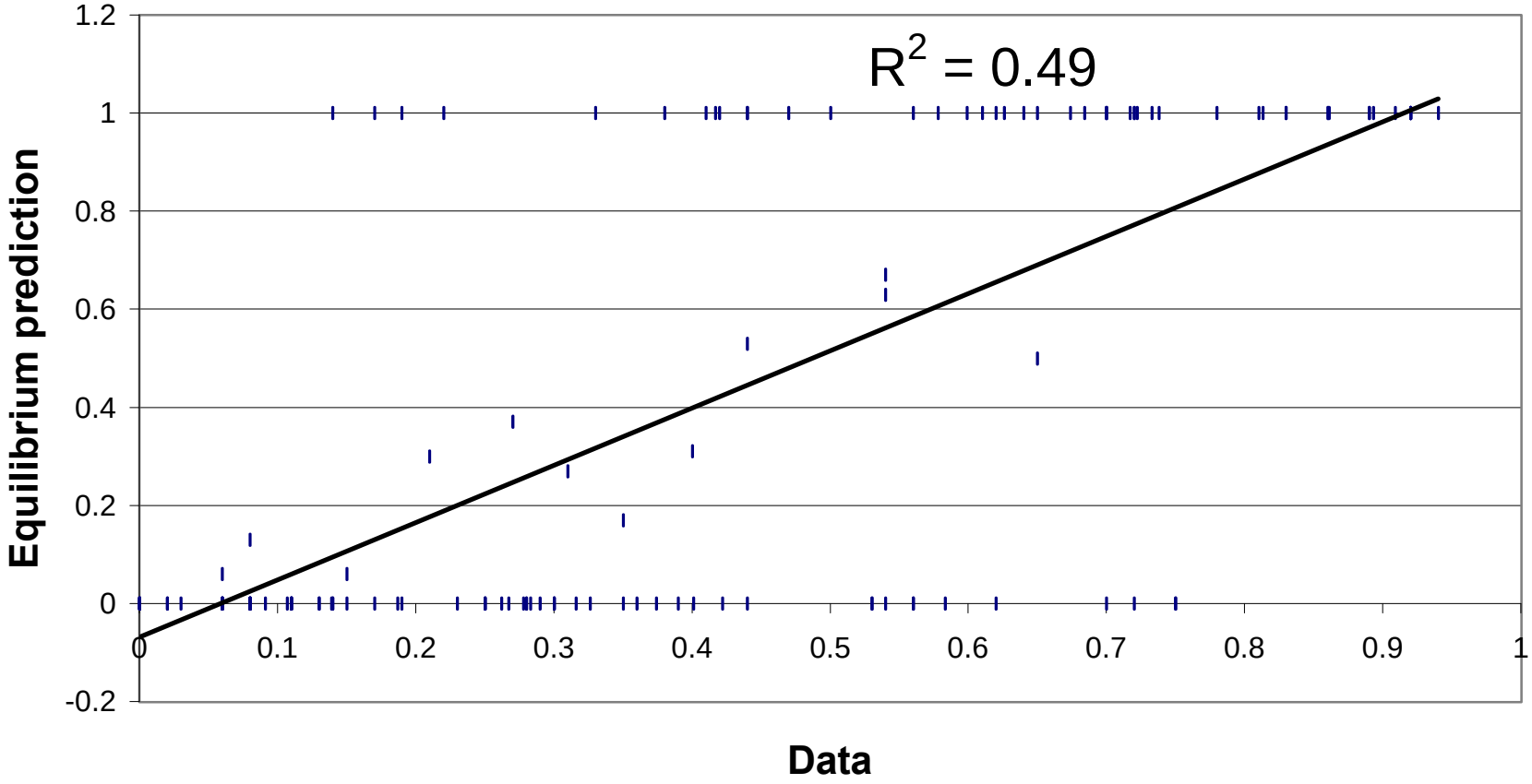


Figure 4: How entry varies with capacity (c), data and thinking-steps model

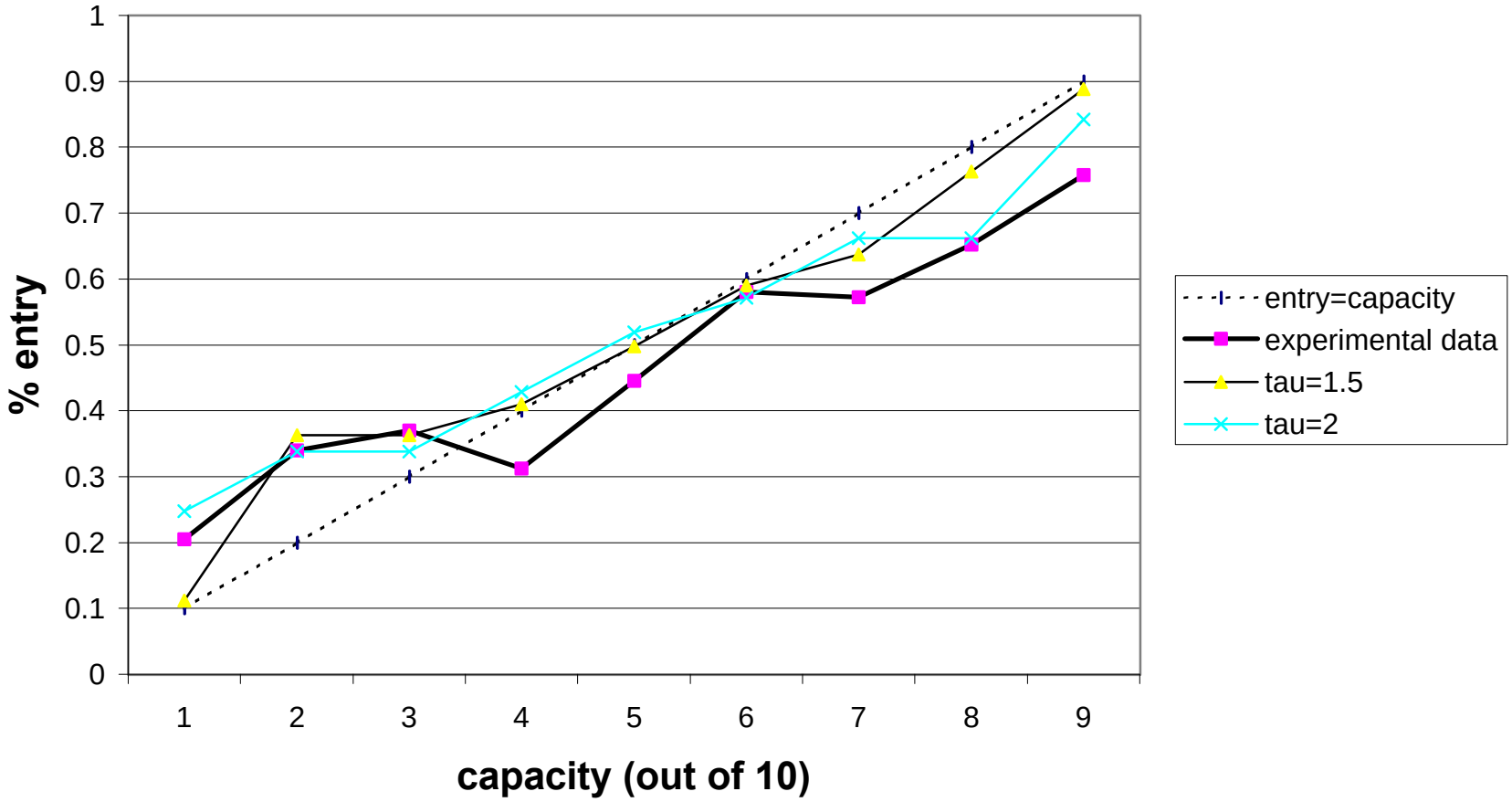
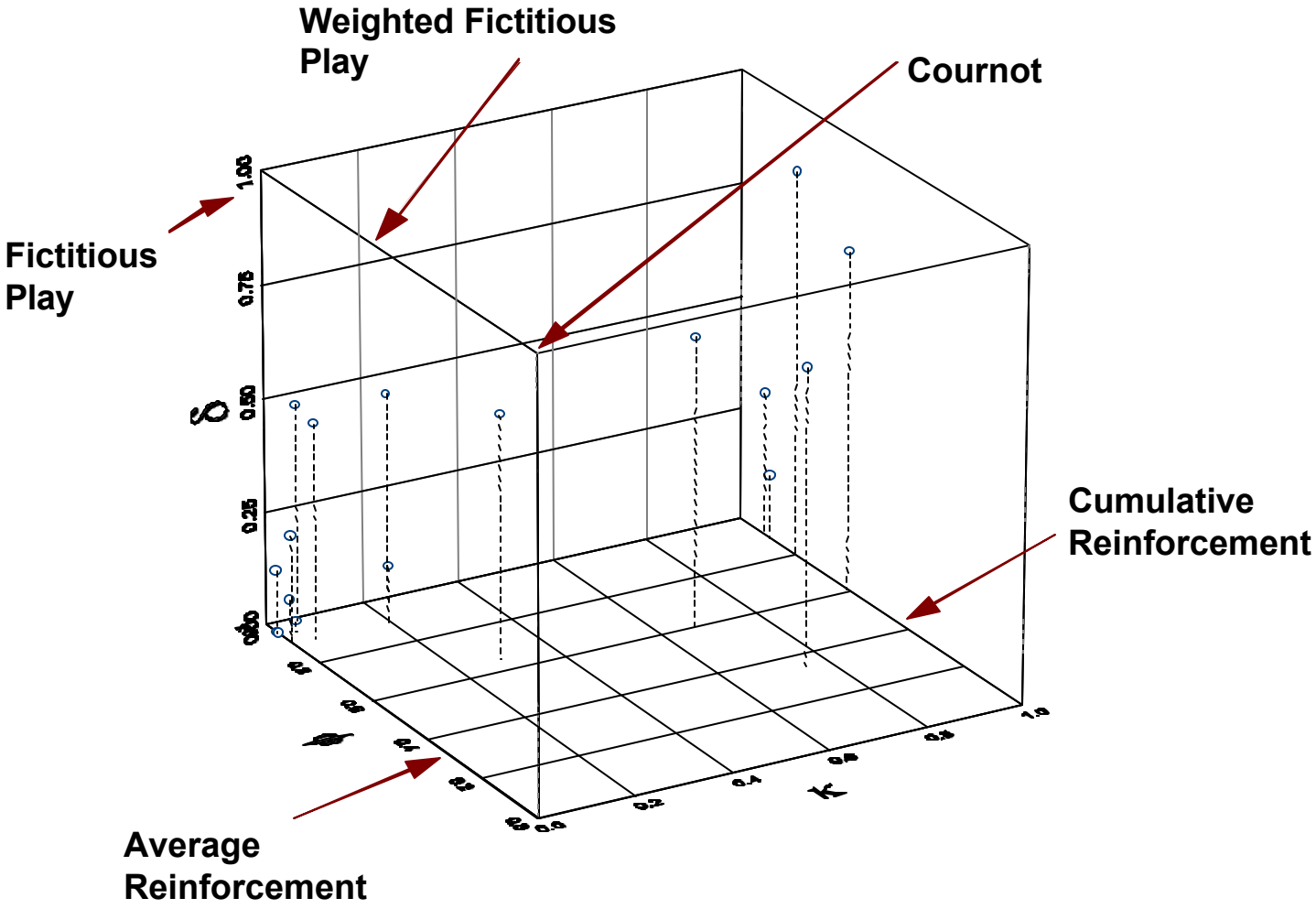


Figure 5: The EWA Learning Cube



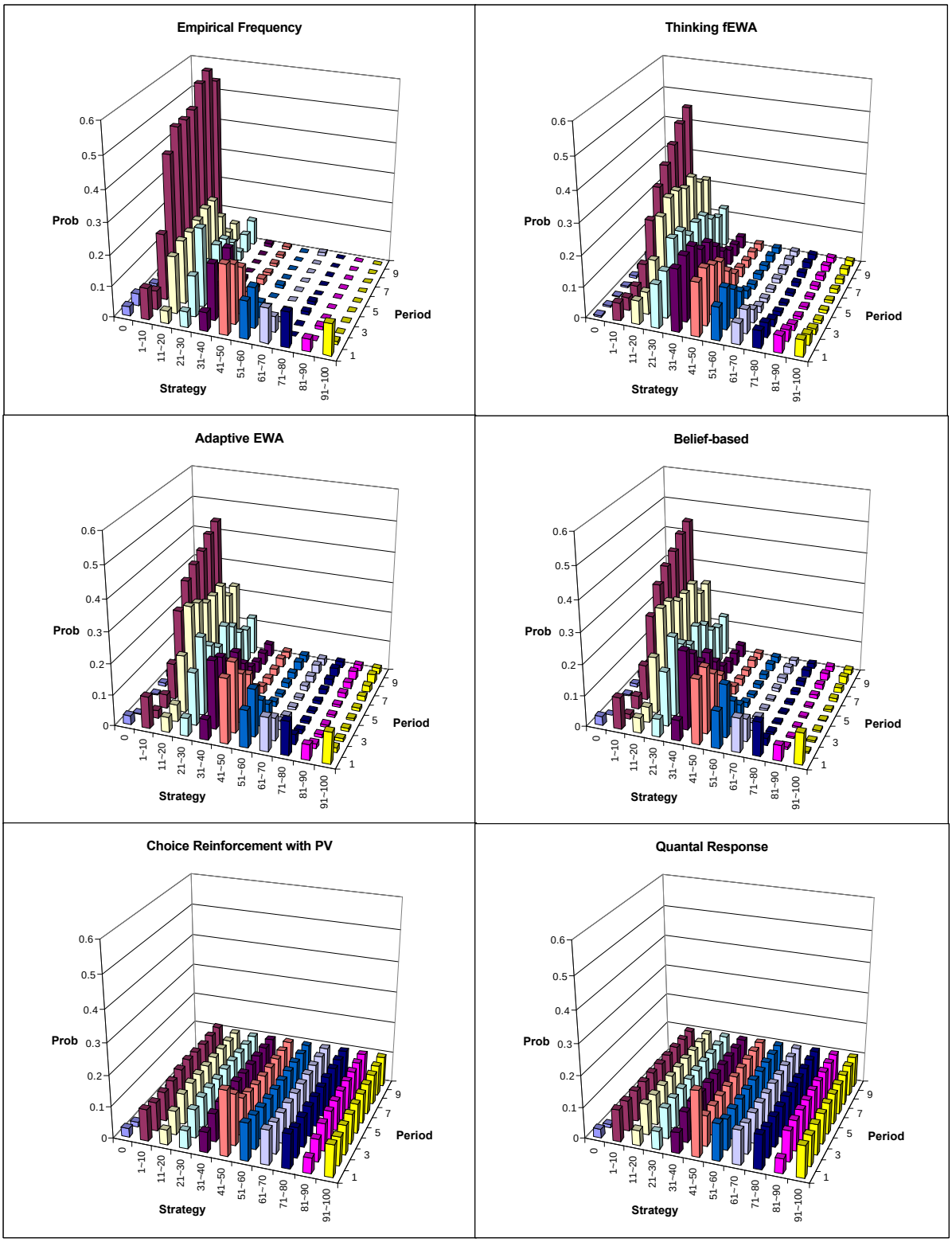


Figure 6: Predicted Frequencies for p-Beauty Contest

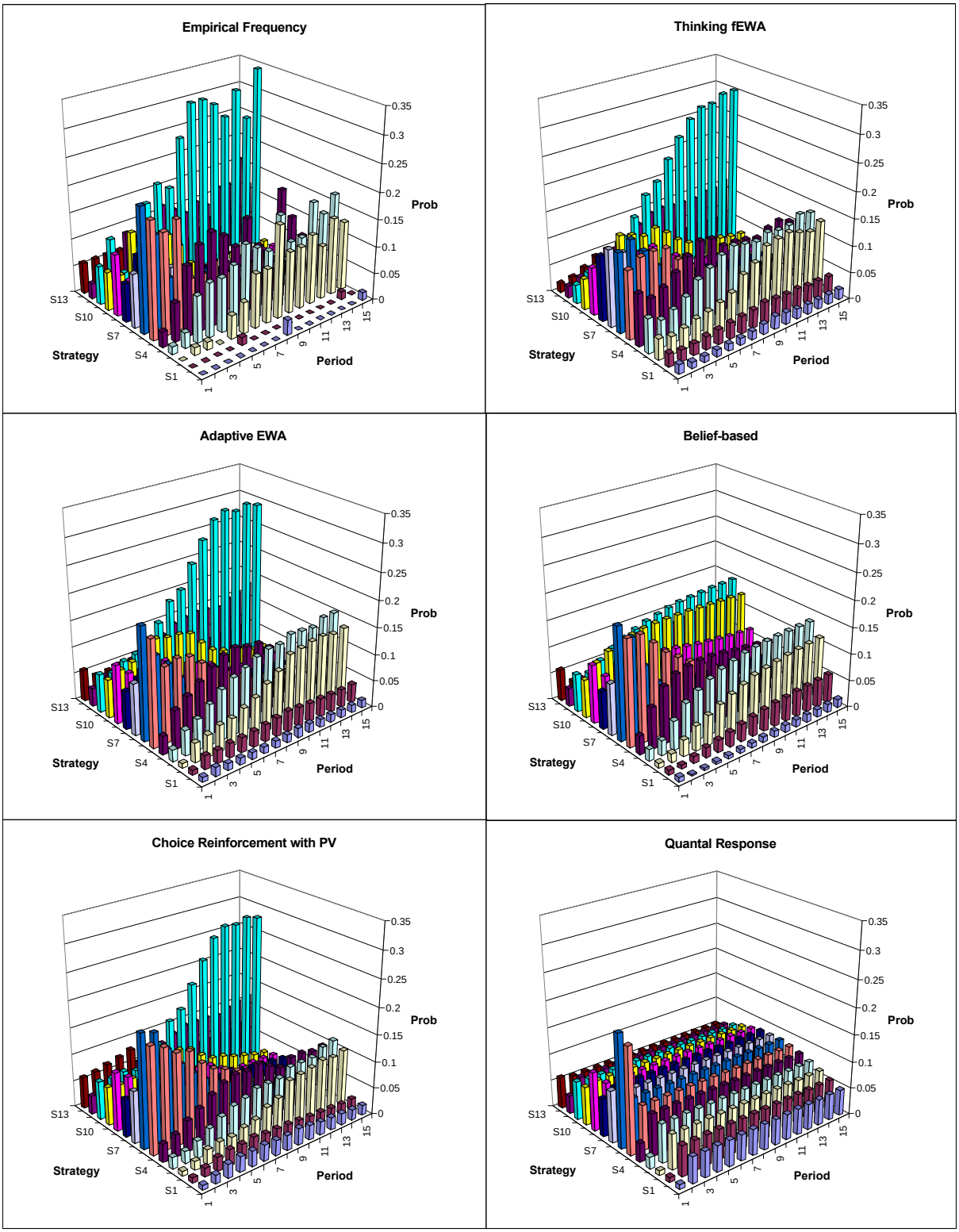


Figure 7: Predicted Frequencies for Continental Divide

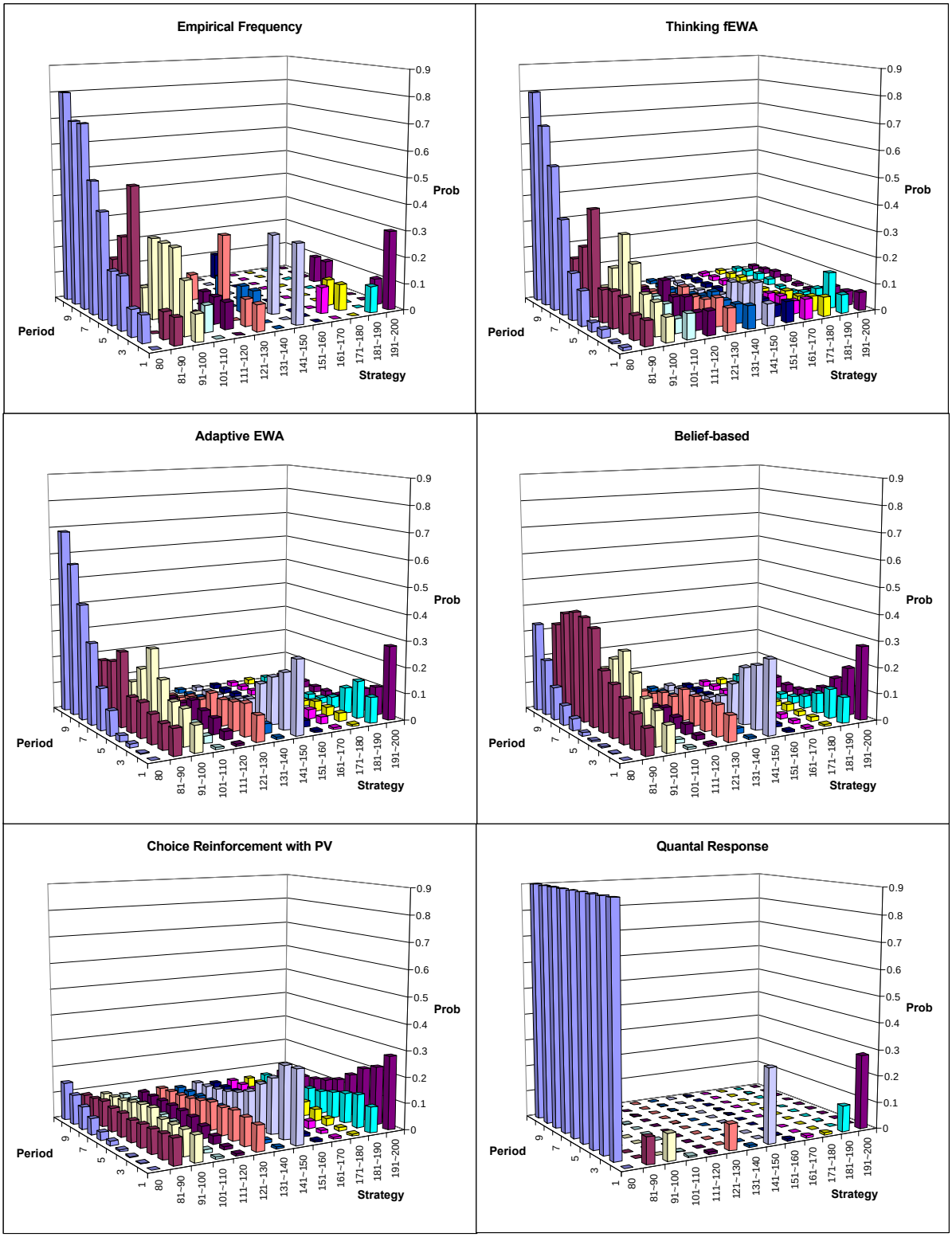


Figure 8: Predicted Frequency for Traveller's Dilemma (Reward = 50)

Figure 9a: Empirical Frequency for No Loan

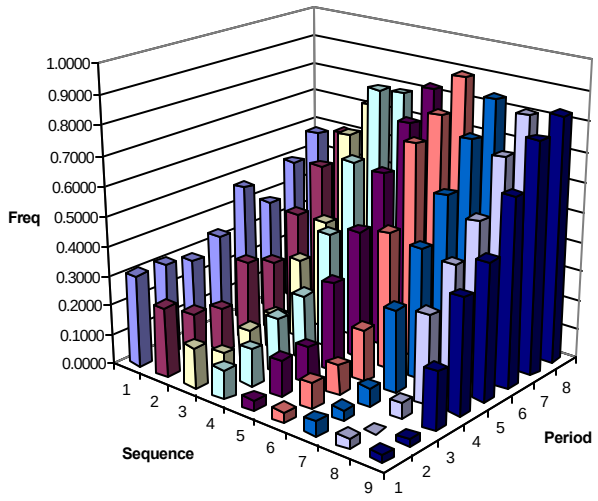


Figure 9c: Predicted Frequency for No Loan

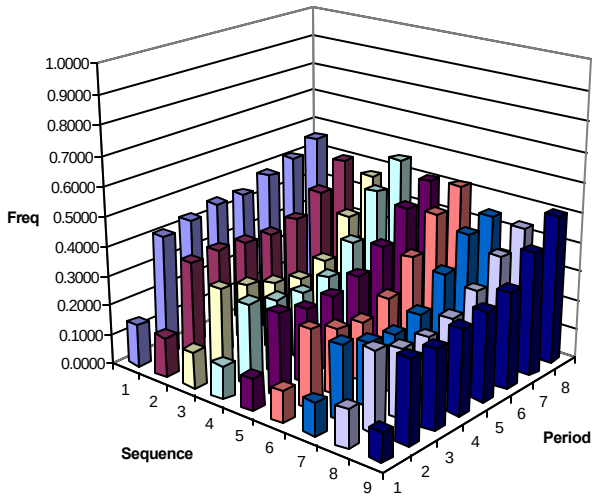


Figure 9b: Empirical Frequency for Default conditional on Loan (Dishonest Borrower)

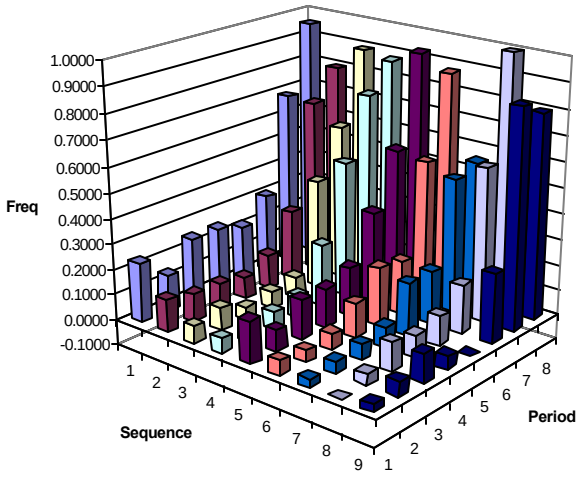


Figure 9d: Predicted Frequency for Default conditional on Loan (Dishonest Borrower)

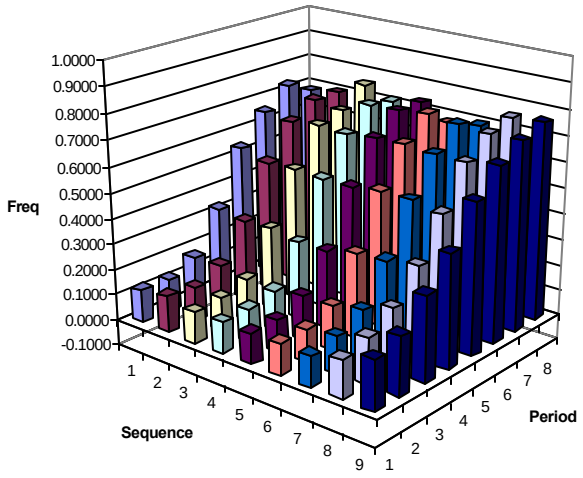


Figure 10: Fit of data to equilibrium & thinking steps model (common t=1.5) in mixed-equilibrium games